

Contents

Preface	viii
1 Introduction	1
Introduction. Reliability Planning. Definition of System Reliability. Overall Target and Allocation to Subsystems. Reliability Modelling and Evaluation. Testing and Data Collection. Evaluation of Alternative Designs. Reliability Report. Scope of This Book.	
2 The Preliminaries	7
Introduction. Sample Space. Events. Random Variables. Probability Laws. Expectation. Variance. Covariance. Moments. Coefficients of Skewness and Excess. Transform Methods. Some Special Distributions (Exponential, Normal, Log-normal, Weibull, Gamma). Stochastic Processes. Probability Distributions. Markov Chains. Equilibrium Distribution. Time Specific Behaviour. First Passage Times. Alternative Approach to First Passage Times. Continuous Parameter Markov Chains. Transient Behaviour. Equilibrium Probability Distribution. First Passage Times. Exercises. References.	
3 Frequency and Associated Concepts	63
Introduction. Interstate Transition Rate. The Concept of Frequency. Time Specific Domain. Methods of Calculation. Steady State Domain. Time Specific Probabilities. Steady State Probabilities. Alternative Interpretation of Mean Cycle Time, Mean Duration and Mean Frequency. The Relationship to Average Values. The Concept of Equivalent Transition Rate. References.	
4 System Reliability	89
Introduction. Definition and Description of the System and its Requirements. Failure Modes and Effects Analysis. State Space Approach. Series System. Parallel Systems. Decomposition Using the Conditional Probability Approach. Network Approach. Network Reduction Procedure. Cut Set or Tie Set Methods. Tie Set Manipulation. Cut Set Manipulation. Frequency Calculation Using the Cut Set Approach. Algorithm to Determine Minimal Cut Sets. Exercises. References.	

5 Techniques for Large Systems	132
Introduction. The Problem Areas. Equivalent Transition Rate and Conditions of Mergeability. Components Subject to Fluctuating Environment. State Space Truncation. Sequential Truncation. References.	
6 Reliability Modelling in Non-Markovian Systems	164
Introduction. The Difficulty with Non-Markovian Processes. Method of Supplementary Variables. Semi-Markov Processes. Device of Stages. References.	
7 Simulation	211
Introduction. Basic Procedure. Random Number Generation. Simulation Model. Timing Controls. Random Sampling. Estimating Reliability Measures. Equilibrium Conditions and Sample Size. Variance Reducing Techniques. References.	
8 Conclusions	225
9 Appendices	227
Appendix I	227
Solution of Simultaneous Linear Equations	
Appendix II	230
Shape of the Hazard Rate Function of Two Series Stage Combinations in Parallel	
Appendix III	233
Hazard Rate Shape of Series Stages in Series with a Distinctive Stage	
Appendix IV	234
Series Stages in Series with Two Parallel Stages	
Appendix V	239
Moments of Stage Combinations	
Appendix VI	242
Calculation of the Jacobian Matrix	
Index	245

Preface

The general area of reliability engineering is extremely wide and in fact encompasses all aspects of engineering technology. Conventional intuitive approaches to the evaluation of system adequacy are not sufficient in modern engineering applications and are gradually being replaced by consistent quantitative techniques. A basic and common requirement in any quantitative procedure is the development of a suitable mathematical model to describe the system. The model may be relatively simple or extremely complex and should be capable of numerical manipulation. This book is devoted entirely to this area and deals with the concepts, philosophy and techniques for reliability model building and evaluation.

The book begins by outlining the elements of reliability planning and discusses the role of modelling in the reliability program plan. Chapter 2 reviews the basic probability theory required in subsequent chapters with emphasis on utilization in system reliability modelling and evaluation. The reader will find that some previous background in probability mathematics is helpful but not necessary. Chapter 3 is the key chapter in the book and discusses the concepts of the frequency balancing approach. These concepts have been used with considerable success in the reliability analysis of repairable systems. This chapter emphasizes the calculation of several measures of system reliability. Chapter 4 is concerned with determining the system reliability characteristics from the statistical information available on the failure and repair cycles of the constituent components. Non-maintained systems have been discussed extensively in the available literature and, therefore, this book is directed towards maintained systems. The theory and procedures are, however, quite general and can be equally applied to non-maintained systems. Chapter 5 is devoted to the utilization of these concepts in the reliability analysis of relatively large systems and some possible problem areas and solutions are presented. The available books generally assume constant transition rates and give a cursory treatment to non-Markovian models. Chapter 6 is devoted to non-Markovian modelling with special emphasis on the device of stages which is a very practical approach. The emphasis in the book is on direct analytical methods. A discussion of system reliability modelling would, however, be incomplete without a discussion on simulation methods as given in Chapter 7.

The scope of application of the concepts outlined in the book encompasses virtually all engineering disciplines. The book is therefore intended as a general treatise and is not aimed at any specific area of application.

C. Singh
R. Billinton

CHAPTER 1

Introduction

Introduction

System designers have always been concerned with the subject of reliability. The general approach has been, however, either intuitive or based on rule of thumb criteria derived from previous experience with similar systems. The intuitive approach has proved to be inadequate with the growth of complex military and industrial systems, where a composite of equipment, skills and techniques function as a unified entity. There has been considerable emphasis, in the past two decades, on the development of quantitative techniques and indices which respond meaningfully to the factors which actually affect the system reliability. Quantitative evaluation is achieved by building mathematical models which reasonably idealize the actual system and can be manipulated to obtain suitable measures of reliability. The role of reliability modelling and evaluation can be appreciated by examining the various stages of a general reliability program.

Reliability Planning

It is generally agreed that system reliability must be built in at the design stage of a project. The desired level of reliability can be achieved only by planning and implementing a good reliability program. A reliability program generally consists of the following elements:

- 1 definition of system reliability
- 2 overall target and allocation to subsystems
- 3 reliability modelling and evaluation
- 4 testing and data collection
- 5 evaluation of alternative designs
- 6 reliability report

This sequence is not rigid and in fact many steps may have to be repeated in an iterative fashion. A brief description of these steps is as follows.

1 Definition of System Reliability

There are several definitions of reliability quoted in the literature but the one most often stated in textbooks is 'the probability that the system will perform

its intended function for a given period of time under stated environmental conditions'. This definition is, however, inadequate for many occasions and is restrictive in its scope of application. It is more appropriate to talk of quantitative measures which when compared with reference indices, indicate expected consistency with or deviation from the required performance. Several measures are discussed in detail in Chapter 3 and a brief review of these is given here. The measures may be time specific, i.e., functions of time, or steady state when they refer to the equilibrium conditions. The former are required when the analyst is concerned with the transient behaviour of the system and the latter while considering the average behaviour over a long time.

It is usual in the literature to define reliability indices in terms of system success or failure. Many complex systems have, however, several levels of failure. For example, a large piece of complex equipment may not be simply working or not working but may have many possible output states. It is therefore appropriate to define the calculated reliability measures in terms of a subset X^+ which may contain any number of system states. In particular applications, X^+ can be referred to as success, failure or some other appropriate name.

Time Specific Domain

The following indices are commonly used for repairable systems in the transient domain:

(a) Time Specific Availability of Subset X^+

This is also called point wise availability or instant availability and is the probability of the system being in any state contained in X^+ at a particular instant of time t .

(b) Fractional Duration of Subset X^+

Also known as the interval availability, fractional duration of X^+ , and is defined as the expected proportion of the interval (t_1, t_2) spent in X^+ .

(c) Interval Frequency

The interval frequency is defined as the expected or mean number of times the subset X^+ is encountered in the interval (t_1, t_2) .

Reliability

If the success and failure states are denoted by X^+ and X^- , then reliability is the probability of being in X^+ at the t without having entered X^- . The term reliability is used in many ways and most often in a qualitative sense to indicate concern regarding the ability of the system to perform its intended function. This approach is extended to qualitative appraisal where the term reliability is

considered as an intrinsic system parameter which can be measured by various indices. The definition given above is more specific and reliability is considered as a mathematical quantity which is itself a measure.

Steady State Domain

(a) Steady State Availability of X^+

Commonly called availability, this is the limiting value of both point wise availability and fractional duration. This can, therefore, be interpreted in two ways. The first is as the probability of being in a state contained in X^+ at some point of time remote from the origin. The second is as the time spent in X^+ as a fraction of the total time $(0, T)$ as T tends to be very large.

(b) Steady State Frequency of Encountering X^+

It is more often simply called frequency and can be defined in two ways. The first, is the mean rate at which X^+ is being encountered at some point in time remote from the origin. The second, is the average number of encounters of X^+ , considered over a very large time interval.

(c) Mean Cycle Time

This is defined as the mean time between two successive encounters of X^+ . It is the reciprocal of frequency.

(d) Mean Duration of X^+

It is the expected time of residence in X^+ in one cycle of X^+ .

In addition to the measures defined above, two more useful measures can be calculated using the concepts outlined in Chapter 2.

(e) Mean First Passage Time

This is the time from system initiation to the first encounter of X^+ where X^+ denotes the system failure condition.

(f) Mean Passage Time

When considered in terms of system failure, the mean passage time is called the mean time to failure and is the mean time from an instant when the system is in X^- (X^- is the disjoint of X^+), chosen randomly, to the encounter of X^+ .

The choice of a proper measure depends upon several factors such as the system requirements, the feasibility of calculations, measurability, and so on. The various measures respond in different ways to the system parameters and no single measure can give a complete picture of system reliability. When dealing with repairable systems in the time domain, the time specific availability, fractional duration and interval frequency are the most useful parameters. In the steady state, availability, frequency or cycle time and the mean duration provide

a good measure of system adequacy. The use of multiple indices can sometimes create problems regarding decision making. A weighting or majority voting procedure can be employed in such cases. Usually one index is more important in a given physical environment due to the nature of the system function.

2 Overall Target and Allocation to Subsystems

The overall reliability target for a given system is normally determined by consultation between management, planning and design functions within the organization. The selected target is based on the state of the art and the need and desire for further improvements. Target selection consists of setting a specific probability, frequency or some other reliability indices as the goals for the project reliability program. In general, the targets should be optimistic but should also be physically achievable. Reliability is like other parameters such as speed, weight and cost and as such it is subject to trade-off with these other parameters. For example, the reliability of a particular pumping system may be greatly improved by installing spare pumps, but the cost may be prohibitive. Once an overall target for the system has been defined, the next step is to allocate the targets for the different subsystems which make up the total system. This can involve a considerable amount of effort and in this regard, the following techniques are helpful.

Similar Familiar System Technique

This approach is based on experience and objective judgement and is very useful in the conceptual design stage prior to total definition of the system. The allocation is based on the state of the art of similar systems performing the same basic function. When using this technique, the following factors must be carefully considered when making the projection:

- (i) system physical and performance comparison
- (ii) design similarity
- (iii) manufacturing similarity

All assumptions and conditions required to meet the predicted target figure and their implications to the program must be examined and defined.

Factors of Influence Method

This technique is used when there is an overall reliability goal and the system or equipment design is new, or when the system is a modification of an existing system for which operating experience data are not available:

- (i) for each system considered, assigning weights to the following factors:

- (a) complexity/time of operation
- (b) environmental conditions
- (c) state of the art
- (d) criticality

- (ii) for each system, adding the above weights to obtain the system weight
- (iii) obtaining the relative system weight for each system by normalization
- (iv) apportioning the overall targets according to the relative system weights.

3 Reliability Modelling and Evaluation

This is an important element in any reliability program because the selected model provides the basis for predicting the reliability measures. Various techniques of reliability modelling and evaluation are discussed in detail in this book. These techniques are either direct analytical modelling or simulation or a mixture of the two approaches. In the direct analytical modelling method a model is built which reasonably idealizes the physical system and is also amenable to calculation. The reliability measures are then obtained by manipulating the model. This approach is superior to simulation and should be used wherever possible. Simulation also employs a mathematical model but proceeds by performing sampling experiments on this model. It is more flexible but is also more time-consuming and less accurate. Simulation can be used to provide estimates of the same basic measures which would be obtained by a direct analytical approach.

In the initial stages of design, when only a general system outline is available, the analyst may have to be satisfied with rudimentary failure modes, effects and criticality analysis (FMECA). This technique can be used to systematically study the modes of failure and their effects on the system. The various failure modes can be arranged in order of their criticality to the system requirements and provide some very useful design modification data. The reliability model is in general modified and improved as the design progresses and becomes solidified.

There is one pitfall which every reliability engineer must guard against. Repeating experiments with a mathematical model on a computer can generate confidence in the reliability measures so obtained. A closer look at the input data and at the model used may prove that this sense of confidence is not really justified. The assumptions built into any model and the validity of the data used must be carefully considered when interpreting the results provided by the model.

4 Testing and Data Collection

In many research and development projects, tests will be made at component, assembly and system level. A systematic procedure should be adopted for collecting data from these trials in a cumulative manner. In many projects it

6 System Reliability Modelling and Evaluation

may not be possible to conduct tests, and data from other sources having similar units may be used. Data collection activities are a vital part of reliability evaluation since without valid data, excessive sophistication in model building may be simply an intellectual exercise.

5 Evaluation of Alternative Designs

Theoretically, a number of designs should be prepared and the one having maximum reliability and satisfying the other constraints should be selected. In some cases, the number of alternatives may be relatively small due to physical constraints. Reliability analysis provides an additional degree of consistency to the evaluation of alternate proposals. In many cases, the decision is based upon lowest total cost and therefore the function of the reliability analyst is to ensure that the system or design selected does satisfy the required targets. Reliability evaluation should provide useful input at the design decision points. If the analysis is not done as the design progresses, it does not provide this initial function and therefore becomes a bookkeeping exercise after the fact.

6 Reliability Report

The final step in any reliability study should be the preparation of a detailed report containing information on the reliability program, the trials made and the results obtained. The report must contain the assumptions made in developing the reliability model and also indicate the level of confidence in the data used. The report should be an objective assessment and should enable the top management to obtain a proper appraisal of the expected system reliability.

Scope of This Book

The general application of reliability engineering concepts is an extremely wide field as illustrated by the six broad topics noted earlier. This book makes no attempt to deal with all these topics and is devoted entirely to the concepts, philosophy and techniques for reliability modelling and evaluation. The systems analyst and the reliability expert must be capable of developing valid models for reliability evaluation of a system. The necessary background in probability and stochastic processes is therefore essential and is reviewed in some detail in Chapter 2. The book emphasizes the calculation of more than one reliability measure as a single index may not provide a complete picture of the system reliability. The book covers both constant transition rate and non-Markovian systems. A special chapter is devoted to the problems encountered in reliability evaluation of large systems and several solutions are proposed. Reliability analysis is an integral part of economic system design. It is extremely important that reliability be considered in quantitative rather than qualitative terms and therefore provide a consistent and responsive indicator of system adequacy.

CHAPTER 2

The Preliminaries

Introduction

An appreciation of certain basic probability concepts is essential for the development of reliability models and their subsequent evaluation. In general, probability mathematics provide the medium for examination of systems which exhibit random phenomena, i.e. behave in accordance with probabilistic rather than deterministic laws. It has been assumed that the reader is familiar with the basic probability concepts normally encountered in an undergraduate engineering course. Many text books and, therefore, courses introduce probability theory either as an abstract mathematical concept or through the use of *a priori* situations such as dice or playing-card type examples. In these cases, it is often difficult to develop an easy and effective interface between the basic reliability and probability concepts. This chapter reviews the basic probability theory required in subsequent chapters with emphasis on utilization in system reliability modelling and evaluation.

Sample Space

The set of all possible outcomes of a random phenomenon is called the sample space, sample description space or possibility space. As an example consider two transmission links each existing either in the up state (U) or the down state (D). The description of the possible states at any time is given by the set

$$S = \{(1U, 2U), (1U, 2D), (1D, 2U), (1D, 2D)\}$$

A set which has a definite number of elements is called a finite set. A set which is not finite but whose elements can be put in a one to one correspondence with the set of natural numbers is said to be countably infinite or denumerable. Both of these types, the finite and the denumerable set come under the general name of countable set.

As another example, consider the load on a pumping system. This may assume any value between L_0 , the minimum load and L_1 the maximum peak load. The sample space S , therefore, consists of all points s such that

$$L_0 \leq s \leq L_1$$

The interval (L_0, L_1) contains a noncountable infinity of members. Such a sample space which is not countable is called uncountably infinite or simply uncountable.

Events

Consider the descriptions $(1U, 2D)$ and $(1D, 2U)$ in the example of two transmission links. These descriptions define the event that one of the transmission links has failed. Similarly if both the transmission links are needed to keep the system in operation, the set $\{(1D, 2D), (1U, 2D), (1D, 2U)\}$ defines the event that the system has failed. An event can therefore be defined as a set of descriptions and the event E is said to have occurred if the outcome of the random phenomenon is a member of E . As the sample space contains the descriptions of all possible outcomes of the random phenomenon under consideration, an event may also be defined as any subset of the sample description space. In the case of demand on the pumping system, the subset $A = \{s: L_2 \leq s \leq L_1\}$ defines the event that load is greater than or equal to L_2 . Whenever the load is in the interval (L_2, L_1) , the event 'load is greater than L_2 ' is said to have occurred. The events are sets and therefore the algebra of events is essentially the concepts of set theory.

Random Variables

A random variable is a quantity which assumes values in accordance with certain probabilistic laws. A random variable which assumes discrete values is called a discrete random variable and one which assumes values from a continuous interval is termed a continuous random variable. This definition of a random variable is sufficient for the purpose of this book. A more precise definition of a random variable is as a function defined on a sample space in which certain technical conditions are satisfied. The random variable relabels the descriptions of outcomes contained in S in terms of real numbers. The domain of the random variable is the sample space S and the range is contained in the set Re of all real numbers.

Consider again the example of the two transmission links. Let X be the function defined on the sample space S , where X denotes the number of components down. The values of this discrete random variable are given below:

$$\begin{aligned} X(s) &= 0, s = (1U, 2U) \\ &= 1, s = (1U, 2D), (1D, 2U) \\ &= 2, s = (1D, 2D) \end{aligned}$$

In the above example, instead of describing the outcomes as shown above, they could be assigned integer values by a random variable Z such that

$$\begin{aligned} Z &= 0, s = (1U, 2U) \\ &= 1, s = (1U, 2D) \\ &= 2, s = (1D, 2U) \\ &= 3, s = (1D, 2D) \end{aligned}$$

The random variable Z depicts the state of the system and its different values are termed the state space. If a device is put into operation at time $t = 0$, the time to failure can be denoted by the continuous random variable X . Some other examples of a continuous random variable X assuming a value x are temperature $(-273^\circ \text{C} \leq x < \infty)$, the load on a power system $(L_0 \leq x \leq L_1)$, the repair time of a component $(0 < x < \infty)$ and the noise voltage at an amplifier output point.

Probability Laws

In application, the functional form of the random variable is not usually of great interest. The main focus is usually on the probability with which the random variable assumes a certain value. Probability is a function which assigns a number between 0 and 1 to a set of points (event) in S . The statement that the random variable X assumes a value in set B of real numbers, implies that the event defined by the subset, $\{s: X(s) \text{ is in } B\}$ occurs. Therefore

$$P[\{s: X(s) \text{ is in } B\}] = P[X \text{ is in } B]$$

This is the basic formula for obtaining the probability function of the random variable X from the probability function which exists on the sample space S on which the random variable X is defined as a function. Some other types of events in terms of a random variable can be defined for fixed numbers x, a, b

$$\begin{aligned} [X = x] &= \{s: X(s) = x\} \\ [X \leq x] &= \{s: X(s) \leq x\} \\ [X > x] &= \{s: X(s) > x\} \\ [a < X < b] &= \{s: a < X(s) < b\} \end{aligned}$$

A discrete random variable assumes only discrete values $x_i, i = 0, 1, 2, \dots$ from

the set $\mathcal{R}$ of real numbers. The probability density function for a discrete random variable is defined by

$$p_X(x) = P[X = x] \quad (2.1)$$

This function should clearly have the following properties

(i) $p_X(x) = 0$ unless x is one of $x_0, x_1, x_2, \dots$

(ii) $0 \leq p_X(x_i) \leq 1$

(iii) $\sum_i p_X(x_i) = \sum_i P[X = x_i] = 1$

The probability density function for a discrete random variable is sometimes also called the probability mass function of X . The probability function $P_X(B)$ of the random variable in terms of probability mass function is given by

$$P_X(B) = P[X \text{ is in } B] = \sum_{x_i \in B} p_X(x_i)$$

The probability distribution function $F_X(x)$ of the discrete random variable is given by

$$\begin{aligned} F_X(x) &= P[X \leq x] \\ &= \sum_{x_i \leq x} p_X(x_i) \end{aligned}$$

It should be noted that the domain of the distribution function is the set of all real numbers and the range, being a probability is the interval (0,1). This is in contrast to the case of a random variable whose domain is the sample space.

A continuous random variable can assume any value over a continuous interval. Since the number of elementary events in a continuous interval is infinite, the probability of the random variable X assuming a value exactly equal to x is zero. This is, however, not an impossible event. For example, although the time to actual failure of a component may be x , the probability of this event happening is zero. The probability density function of the form in Equation (2.1) is therefore not suitable for a continuous random variable. The probability density function $f_X(x)$ of a continuous random variable X is so defined that

$$P[a \leq X \leq b] = \int_a^b f_X(y) dy \quad (2.2)$$

The probability distribution function $F_X(x)$ can be now written as

$$\begin{aligned} F_X(x) &= P[-\infty < X \leq x] \\ &= \int_{-\infty}^x f_X(y) dy \end{aligned} \quad (2.3)$$

It follows from Equation (2.3) that

$$f_X(x) = F'_X(x) \quad (2.4)$$

The function $f_X(x)$ has the following properties

1. $f_X(x)$ is non-negative

$$2. \int_{-\infty}^{+\infty} f(x) dx = 1$$

3. The function $f_X(x)$ is continuous at all but a finite number of points, i.e. it is piece-wise continuous.

When we are dealing with the distributions of operating and down times, the random variable X is non-negative and its density function is zero over the negative range. It is sometimes more convenient to work with the complementary function of $F(x)$ called the survivor function

$$\begin{aligned} \mathcal{F}(x) &= P[X > x] \\ &= \int_x^{\infty} f(y) dy \\ &= 1 - F(x) \end{aligned} \quad (2.5)$$

It follows from this expression that

$$f(x) = -\mathcal{F}'(x) \quad (2.6)$$

In reliability modelling and evaluation one function which is used extensively and which is equivalent to $f(x)$ is the hazard function. In practice, depending upon the circumstances employed, it may be known by a variety of names, age specific failure rate or simply failure rate, repair rate, hazard rate, force of mortality etc. A detailed interpretation of this function is given in the next

chapter while developing the concept of frequency. This can be defined as

$$\phi(x) = \lim_{\Delta x \rightarrow 0^+} \frac{P[x < X \leq x + \Delta x | x < X]}{\Delta x} \quad (2.7)$$

With the interval length approaching zero, Equation (2.2) can be written as

$$f_X(x) = \lim_{\Delta x \rightarrow 0^+} \frac{P[x < X \leq x + \Delta x]}{\Delta x}$$

Equation (2.7) can be rewritten in the form

$$\begin{aligned} \phi(x) &= \lim_{\Delta x \rightarrow 0^+} \frac{P[x < X \leq x + \Delta x]}{\Delta x} \cdot \frac{1}{P[x < X]} \\ &= \frac{f_X(x)}{\mathfrak{F}_X(x)} \end{aligned} \quad (2.8)$$

It is usual to drop the suffix X denoting the random variable unless there is a likelihood of incorrect interpretation and this has been done in this text. Using Equations (2.6) and (2.8)

$$\begin{aligned} \phi(x) &= -\frac{\mathfrak{F}'(x)}{\mathfrak{F}(x)} \\ &= -\frac{d}{dx} [\log \mathfrak{F}(x)] \end{aligned} \quad (2.9)$$

Integrating Equation (2.9) and using the condition that $\mathfrak{F}(0) = 1$

$$\mathfrak{F}(x) = \exp \left[-\int_0^x \phi(y) dy \right]$$

and therefore

$$f(x) = \phi(x) \exp \left[-\int_0^x \phi(y) dy \right]$$

This expression shows that $\phi(x)$, the hazard function uniquely determines the probability density function. The probability density function, the probability distribution function, the survivor function and the hazard rate function are mathematically equivalent.

Expectation

The probabilistic behaviour of a random variable is completely defined by the probability density function or distribution function. It is, however, often of interest to obtain a single value which may represent the random variable and its probability distribution. One such characteristic value is the expectation or mean. This is denoted by $E(x)$ and given by

$$E(X) = \sum_i x_i P[X = x_i]$$

$$= \sum_i x_i p(x_i) \quad \text{if } X \text{ is a discrete random variable}$$

or

$$= \int_{-\infty}^{+\infty} x f(x) dx \quad \text{if } X \text{ is a continuous random variable} \quad (2.10)$$

The expectation is said to exist if the series or the integral involved converges absolutely, i.e.

$$\sum_i |x_i| p(x_i) < \infty \quad \text{for the discrete case}$$

and

$$\int_{-\infty}^{+\infty} |x| f(x) dx < \infty \quad \text{for the continuous case}$$

The expectation or the mean value has a meaningful interpretation in terms of the average of a sample. This interpretation is provided by the law of large numbers. Assume that there are n random variables $X_1, X_2, \dots, X_n$ which are identically distributed as X and each has a mean m . This set of n random variables represents a random sample of n observations $X_1, X_2, \dots, X_n$. The sample mean, which is also a random variable is given by

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

According to the law of large numbers for any constant $c > 0$

$$\lim_{n \rightarrow \infty} P[|\bar{X} - m| > c] = 0 \quad (2.11)$$

This implies that as the sample size increases, the sample mean approaches the mean of the random variable. Thus if the random variable X is observed many times and each time the arithmetic mean is calculated, it will approach

the mean or the expectation of the random variable X as the number of observations becomes very large.

An important result concerning the sum of the random variables is given by

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$$

That is the expectation of the sum of a group of random variables is equal to the sum of the expectations of the random variables. This result holds even if the random variables are not independent.

Variance

The arithmetic mean indicates the central tendency and provides a value around which the random variable X is distributed. Two or more random variables may have the same mean value but the deviations from this value may have different likelihoods. The information regarding the scatter of the values around the mean is provided by the variance of X , designated $V(X)$

$$V(X) = E[(X - E(X))^2] \quad (2.12)$$

$$= \sum_i (x_i - m)^2 p(x_i) \quad \text{for the discrete case}$$

and

$$= \int_{-\infty}^{+\infty} (x - m)^2 f(x) dx \quad \text{for the continuous case}$$

The variance is a weighted average of the values $(X - m)^2$ and therefore if large it means that the distribution of X is such that large deviations occur with a comparatively high probability. On the other hand if the values are near m with large probability, the variance is comparatively small. The variance, therefore, provides us with a measure of the spread of the distribution. Equation (2.12) can be put into another form more suitable for computation

$$\begin{aligned} V(X) &= E[(X - E(X))^2] \\ &= E(X^2) - (E(X))^2 \end{aligned} \quad (2.13)$$

The variance is often denoted by σ^2 and has the dimensions of X^2 . The square root of the variance σ has the dimensions of X and is called the standard deviation.

Covariance

Consider the two random variables X and Y . The quantities $(X - E(X))$ and $(Y - E(Y))$ are the deviations of the two random variables from their respective means. The expected value of the product of these deviations is called the covariance and is given by

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E(XY) - E(X)E(Y) \end{aligned} \quad (2.14)$$

In terms of the joint probability density function $f(x, y)$ of X and Y , it is given as

$$\text{Cov}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - m_x)(y - m_y) f(x, y) dx dy \quad (2.15)$$

where $f(x, y)$ is so defined that

$$P(x < X \leq x + dx, y < Y \leq y + dy) = f(x, y) dx dy$$

The covariance gives a measure of the tendency of the two random variables to vary together. If both the deviations have the same sign, then the sign of the product is positive and otherwise it is negative. Therefore, if the two variables tend to vary in harmony, the sign of the product will be positive with a larger probability and the covariance will, therefore, be positive. On the other hand if the two variables tend to vary in opposition, the sign of the covariance is negative. When the two random variables are independent, their covariance is zero. This tendency to vary together or in opposition is often measured by the dimensionless quantity called the correlation coefficient

$$C_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)} \sqrt{V(Y)}} \quad (2.16)$$

which can be shown to lie in the range $[-1, 1]$. The correlation coefficient is the covariance normalized by the product of the standard deviations and its sign is the same as that for covariance. The statements made about the sign of the covariance, therefore, apply to the correlation coefficient too. It should be noted that correlation and non-correlation are similar to independence and interdependence but the two are not identical. If the two random variables are independent then $C_{XY} = 0$ but the reverse is not always true, i.e. independence does not necessarily follow from non-correlation. Correlation

implies dependence but the reverse is not necessarily true i.e. $C_{XY} \neq 0$ does not follow from dependence.

Moments

The expectation of a real valued function $g(X)$ of X is given by

$$\begin{aligned} \text{(i)} \quad E[g(X)] &= \sum_i g(x_i) p(x_i) \quad \text{if } X \text{ is discrete} \\ \text{(ii)} \quad E[g(X)] &= \int_{-\infty}^{+\infty} g(x) f(x) dx \quad \text{if } X \text{ is continuous.} \end{aligned} \quad (2.17)$$

One simple function whose expectation is of interest is the k th power of X , i.e. $g(x) = X^k$. The expectation of this function is called the k th initial moment of X or the distribution of X . Therefore the k th initial moment of X is given by

$$\begin{aligned} m_k(X) &= m_k \\ &= \sum_i x_i^k p(x_i) \quad \text{for the discrete case} \\ &= \int_{-\infty}^{+\infty} x^k f(x) dx \quad \text{for the continuous case.} \end{aligned} \quad (2.18)$$

A somewhat more useful concept, the k th central moment, is similarly defined by

$$\begin{aligned} M_k(X) &= M_k \\ &= \sum_i (x_i - m)^k p(x_i) \quad \text{if } X \text{ is discrete} \\ &= \int_{-\infty}^{+\infty} (x - m)^k f(x) dx \quad \text{if } X \text{ is continuous.} \end{aligned} \quad (2.19)$$

It can be seen that the mean is the first initial moment, the mean square value is the second initial moment and the variance is the second central moment.

In a similar manner, the mixed initial moment for the multi-dimensional distribution can be defined as

$$m_{k_1, k_2, \dots, k_n}(X_1, X_2, \dots, X_n) = E(X_1^{k_1} X_2^{k_2} \dots X_n^{k_n}) \quad (2.20)$$

and the mixed central moment is given by

$$\begin{aligned} M_{k_1, k_2, \dots, k_n}(X_1, X_2, \dots, X_n) \\ = E[(X_1 - E(X_1))^{k_1} (X_2 - E(X_2))^{k_2} \dots (X_n - E(X_n))^{k_n}] \end{aligned} \quad (2.21)$$

The covariance is therefore the first central mixed moment of X and Y .

Coefficients of Skewness and Excess

It can be seen from Equation (2.19) that odd moments vanish for a distribution which is symmetrical about the mean value. The third moment can, therefore, be used as a measure of the asymmetry of the distribution. Asymmetry is more conveniently measured as a dimensionless quantity and the asymmetry coefficient or skew coefficient is given by

$$A = M_3/(M_2)^{3/2} \quad (2.22)$$

The coefficient of excess also gives information about the form of distribution by comparing it with the normal probability density function near its mode and is defined by

$$G = M_4/(M_2)^2 - 3 \quad (2.23)$$

$M_4 = 3M_2^2$ for the normal distribution and therefore the coefficient of excess is zero. In the case of a distribution having the same variance as the normal, $G > 0$ indicates that the distribution has a sharper peak than the normal distribution and similarly $G < 0$ indicates a comparatively flatter peak.

Transform Methods

Operational or transform methods are employed for transforming the problem into a functional form which at first glance appears to have nothing to do with the original problem but which often facilitates its solution. They are used in many branches of mathematics especially in solving differential equations. One technique of importance in probability theory is the method of characteristic functions. The characteristic function of a random variable X is defined by

$$\begin{aligned} \phi_X(\theta) &= E[e^{i\theta X}] \\ &= \int_{-\infty}^{+\infty} \exp(i\theta x) f(x) dx \end{aligned} \quad (2.24)$$

where

$$i = \sqrt{-1}$$

and $f(x)$ = the probability density function of X

The probability density function can be found from the characteristic function by the inversion formula

$$f(x) = \int_{-\infty}^{+\infty} \phi(\theta) \exp(-i\theta x) d\theta \quad (2.25)$$

There is a one to one correspondence between the characteristic function and the probability density function. Therefore, two characteristic functions which are equal at all points are the characteristic functions of the same distribution function and vice versa.

Differentiating (2.24) k times

$$\phi^{(k)}(\theta) = \frac{d^k \phi(\theta)}{d\theta^k} = i^k \int_{-\infty}^{+\infty} x^k \exp(i\theta x) f(x) dx$$

This is with the assumption that the k th derivative exists. As $\theta \rightarrow 0$

$$\begin{aligned} \phi^{(k)}(0) &= i^k \int_{-\infty}^{+\infty} x^k f(x) dx \\ &= i^k m_k \end{aligned}$$

Therefore the k th moment of the random variable X , if it exists, can be obtained from the k th derivative of the characteristic function evaluated at zero

$$m_k = i^{-k} \phi^{(k)}(0) \quad (2.26)$$

An important result in probability theory is that the characteristic function of the sum of several independent random variables is the product of the characteristic functions of the random variables. Let

$$Y = X_1 + X_2 + \dots + X_n$$

where X_i are the independent random variables. Then

$$\phi_Y(\theta) = E(e^{i\theta Y}) = E(e^{i\theta(X_1 + X_2 + \dots + X_n)})$$

Since the random variables are independent

$$\begin{aligned} \phi_Y(\theta) &= E(e^{i\theta X_1}) E(e^{i\theta X_2}) \dots E(e^{i\theta X_n}) \\ &= \phi_{X_1}(\theta) \phi_{X_2}(\theta) \dots \phi_{X_n}(\theta) \end{aligned} \quad (2.27)$$

Another transform often used is the moment generating function, defined for all real numbers by

$$\begin{aligned} \psi(\theta) &= E(e^{\theta X}) \\ &= \int_{-\infty}^{+\infty} e^{\theta x} f(x) dx \quad \text{for the continuous case.} \\ &= \sum_i e^{\theta x_i} p(x_i) \quad \text{for the discrete case.} \end{aligned} \quad (2.28)$$

As in the case of the characteristic function, the following results are important

$$m_k = \psi^{(k)}(0) \quad (2.29)$$

and

$$\psi_Y(\theta) = \psi_{X_1}(\theta) \psi_{X_2}(\theta) \dots \psi_{X_n}(\theta) \quad (2.30)$$

when

$$Y = X_1 + X_2 + \dots + X_n$$

X_j being independent random variables.

The Expression (2.29) is suitable for calculating the k th initial moment and the central moments can be calculated from the initial moments. The moment generating function about the mean m or in fact any other point can also be defined as

$$\begin{aligned} \psi_m(\theta) &= E(e^{(X-m)\theta}) \\ &= e^{-m\theta} \psi(\theta) \end{aligned} \quad (2.31)$$

The k th moment about the mean can be calculated from its k th derivative evaluated at zero, i.e.

$$M_k = \psi_m^{(k)}(0) \quad (2.32)$$

The main advantage of the characteristic function over the moment generating function is that it always exists whereas the moment generating function may not exist.

For an integer valued discrete random variable X having a probability mass function p_i , the moment generating function can be simplified by substituting $z = e^\theta$

$$\psi(Z) = \sum_i z^i p_i \quad (2.33)$$

This function is called the probability generating function of X or the Z transform of X . It can be shown that

$$p_k = \frac{\psi^{(k)}(0)}{k!} \quad (2.34)$$

$$m_k = \psi^{(k)}(1) \quad (2.35)$$

Reliability modelling is usually concerned with non-negative random variables and for these the Laplace transform is very useful. In general if $g(t)$ is a real function of a variable t defined for $t > 0$, then the Laplace transform of this function is given by

$$L[g(t)] = \bar{g}(s) = \int_0^{+\infty} g(t) e^{-st} dt \quad (2.36)$$

where s is a complex variable. The inverse Laplace transform is defined by

$$L^{-1}[\bar{g}(s)] = g(t) = \frac{1}{2\pi} \int_{c-i\infty}^{c+i\infty} \bar{g}(s) e^{st} ds$$

when $i = \sqrt{-1}$

In practice it is seldom necessary to perform this contour integration and the inverse is calculated by expanding $\bar{g}(s)$ into partial fractions and using the tables of Laplace transforms. The following results are of importance with respect to this book.

1. The Laplace transform of the derivative of a function $g(t)$ whose Laplace transform is $\bar{g}(s)$ is given by

$$L\left[\frac{dg(t)}{dt}\right] = s\bar{g}(s) - g(0^+) \quad (2.37)$$

Here $g(0^+)$ is the limit of $g(t)$ as t approaches zero from positive values. Very often this will be written simply as $g(0)$.

2. The Laplace transform of an integral is given by

$$L\left[\int_0^t g(u) du\right] = \frac{\bar{g}(s)}{s} \quad (2.38)$$

3. The Initial Value Theorem states

$$g(0^+) = \lim_{t \rightarrow 0} g(t) = \lim_{s \rightarrow \infty} s\bar{g}(s) \quad (2.39)$$

4. The Final Theorem states that

$$g(\infty) = \lim_{t \rightarrow \infty} g(t) = \lim_{s \rightarrow 0} s\bar{g}(s) \quad (2.40)$$

if the limit $g(t)$ exists

While considering the Laplace transform of a probability density function, it should be observed that except for the sign of s it is the same as the moment generating function. Therefore

$$L[f(x)] = \tilde{f}(s) = E(e^{-sX})$$

As $f(x)$ is a probability density function it follows from (2.36) that

$$\tilde{f}(0) = 1$$

Using the result (2.38), the Laplace transform of the probability distribution function is given by

$$\bar{F}(s) = \frac{\tilde{f}(s)}{s} \quad (2.41)$$

and the Laplace transform of the survivor function is given by

$$\bar{\mathcal{F}}(s) = \frac{1 - \tilde{f}(s)}{s} \quad (2.42)$$

The $\tilde{f}(s)$ can be written as

$$\tilde{f}(s) = E\left[\sum_{k=0}^{\infty} (-1)^k \frac{s^k X^k}{k!}\right]$$

$$= \sum_{k=0}^{\infty} (-1)^k \frac{s^k}{k!} m_k \quad (2.43)$$

The k th initial moment of X is given by the coefficient of $\frac{(-s)^k}{k!}$ in the Taylor expansion of $\bar{f}(s)$. Similarly for the k th central moment M_k , $e^{sm}\bar{f}(s)$ should be expanded.

Some Special Distributions

This book is mainly concerned with distributions of continuous random variables such as the time to failure and time to repair. Any function which is non-negative and whose integral over its range equals one, can be a probability density function. There are, however, some special mathematical forms which are often used in reliability modelling as probability density functions. The following section examines these special distributions.

1 Exponential Distribution

A non-negative continuous random variable is said to have a negative exponential distribution if it has a probability density function defined as

$$f(x) = \rho e^{-\rho x} \quad (2.44)$$

where ρ is some positive constant. The corresponding distribution function is given by

$$F(x) = \int_0^x f(u) du = 1 - e^{-\rho x} \quad (2.45)$$

The survivor function is

$$R(x) = \int_x^{\infty} f(u) du = e^{-\rho x} \quad (2.46)$$

The hazard function is

$$\begin{aligned} \phi(x) &= \frac{f(x)}{R(x)} \\ &= \rho \end{aligned} \quad (2.47)$$

The Laplace transform of this distribution is given by

$$\bar{f}(s) = L[\rho e^{-\rho x}] = \frac{\rho}{\rho + s}$$

$$\text{The mean} = (-1) \bar{f}^{(1)}(s) |_{s=0} = \rho(\rho + s)^{-2} |_{s=0} = \frac{1}{\rho} \quad (2.48)$$

It can be seen from Expressions (2.47) and (2.48) that the hazard rate for the exponential distribution is constant and equals the reciprocal of the mean of the random variable having this distribution. Since the hazard rate uniquely determines the probability density function, it follows that if the hazard is constant then

$$\begin{aligned} f(x) &= \rho \exp \left[- \int_0^x \rho du \right] \\ &= \rho e^{-\rho x} \end{aligned}$$

That is the distribution is exponential. It is worth repeating that only the exponential distribution has the property that the hazard rate is constant and is equal to the reciprocal of the mean. The second initial moment is

$$m_2 = (-1)^2 \bar{f}^{(2)}(s) |_{s=0} = \frac{2}{\rho^2}$$

and therefore the variance

$$\begin{aligned} M_2 &= \frac{2}{\rho^2} - \frac{1}{\rho^2} \\ &= \frac{1}{\rho^2} \end{aligned} \quad (2.49)$$

The standard deviation is $\frac{1}{\rho}$ and the coefficient of variation equals one. The graphs of the probability density, probability distribution and the hazard function are shown in Fig. 2.1. The exponential law corresponds to the maximum randomness of the lifetimes of the components. The life of a component observing the exponential law has the interesting property that the previous operating time of the component does not effect the residual or remaining lifetime distribution. Consider that a component whose lifetime is represented by an exponentially distributed random variable X is operated up to

time t then the distribution function of $Y = X - t$ is given by

$$\begin{aligned}
 F_Y(x) &= P[X - t \leq x | X > t] \\
 &= P[X \leq x + t | X > t] \\
 &= \frac{P[t < X \leq x + t]}{P[X > t]} \\
 &= \frac{\int_t^{x+t} \rho e^{-\rho u} du}{e^{-\rho t}} \\
 &= 1 - e^{-\rho x} \\
 &= F_X(x)
 \end{aligned}$$

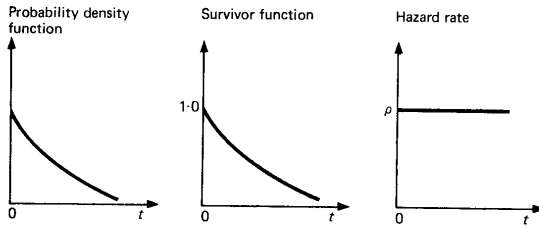


Fig. 2.1 Characteristics of the exponential distribution

It can, therefore, be seen that the distribution of the residual life time of the component is independent of the time for which the component has been operating. It is as if the component forgets how long it has been operating and the breakdown occurs not because of gradual deterioration but a randomly occurring failure. The negative exponential is the only probability distribution having this memory loss property. Its proof is left to the reader in Exercise 1.

The exponential distribution bears an important relationship with the Poisson distribution that if the number of events per unit time is given by a Poisson distribution then the distribution of the interevent time is exponential. Suppose that a component on failure is replaced by an identical component and the number of failures per unit time is λ , then if failure occurs according to the Poisson law, the distribution of failures in time t is given by

$$p_k(t) = P[\text{Number of failures} = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

That no failure occurs during the interval $(0, t)$ is equivalent to having the inter failure time greater than t , i.e.

$$1 - F(t) = p_0(t) = e^{-\lambda t}$$

Differentiating both sides with respect to t

$$f(t) = \lambda e^{-\lambda t}$$

It can therefore be seen that for Poisson failures, the distribution of the time between failures is exponential.

The exponential distribution has desirable mathematical properties and is used widely for representing operating and repair times. The data collected in connection with complex equipment shows that their time of operation is indeed well described by the exponential law. It can be shown mathematically that if a complex piece of equipment has a large number of stochastically independent components such that each component is replaced immediately on failure and every component causes equipment failure, then after a long time the distribution of operating time of the equipment is well approximated by the exponential law. Though the exponential law generally holds for the operating time during the useful life of the equipment, no single hypothesis seems to exist for the repair time. It is generally realized that an exponential representation for the repair time may not be valid in many cases. When the operating time is exponentially distributed and its mean value is much larger than the mean down time, then the steady state values are not generally significantly affected by the nature of the repair time distribution. The question of non-exponential distributions for down times is examined in detail in Chapter 6.

2 Normal Distribution

A continuous random variable is said to be normally distributed if its probability density function is of the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2} dx, \quad -\infty < x < \infty \quad (2.50)$$

where σ is positive and m is any constant. The graph of (2.50) is bell-shaped and is symmetrical about $x=m$ as shown in Fig. 2.2. In the special case when $m = 0$ and $\sigma=1$, this function is called the standard normal density function.

The probability distribution for the normally distributed random variable is given by

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(u-m)^2/2\sigma^2} du \quad (2.60)$$

This function cannot be expressed in closed form in terms of familiar elementary functions. The function $F(x)$ can be transformed into the standard normal distribution by a simple change of variable

$$z = \frac{u-m}{\sigma}$$

Expression (2.60) becomes

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-z^2/2} dz$$

The values of the standard normal probability distribution function are obtained by numerical integration and are tabulated in every elementary statistics book. The use of these tables is left as an exercise to the reader. The moment generating function of a normally distributed random variable is

$$\begin{aligned} \psi(\theta) &= E(e^{\theta X}) \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{x\theta} e^{-(x-m)^2/2\sigma^2} dx \end{aligned}$$

Making the substitution

$$z = \frac{x-m}{\sigma} \text{ so that } x = \sigma z + m \text{ and } dx = \sigma dz$$

$$\begin{aligned} \psi(\theta) &= e^{m\theta} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(z^2-2\sigma\theta z)/2} dz \\ &= e^{m\theta + \sigma^2\theta^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(z-\sigma\theta)^2/2} dz \end{aligned}$$

Making the substitution $w = z - \sigma\theta$, $dz = dw$

$$\begin{aligned} \psi(\theta) &= e^{m\theta + \sigma^2\theta^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-w^2/2} dw \\ &= e^{m\theta + \sigma^2\theta^2/2} \end{aligned} \quad (2.61)$$

$$\begin{aligned} E(X) &= \text{mean} = \psi^{(1)}(\theta)|_{\theta=0} = (m + \sigma^2\theta) e^{m\theta + \sigma^2\theta^2/2}|_{\theta=0} \\ &= m \end{aligned}$$

$$\begin{aligned} E(X^2) &= \psi^{(2)}(\theta)|_{\theta=0} = [(m + \sigma^2\theta)^2 + \sigma^2] e^{m\theta + \sigma^2\theta^2/2}|_{\theta=0} \\ &= m^2 + \sigma^2 \end{aligned}$$

Therefore

$$\text{Var}(X) = E(X^2) - (E(x))^2 = \sigma^2$$

The parameters m and σ , therefore, represent the mean and standard deviation of the normal distribution. The domain of the normal distribution, extends from $-\infty$ to $+\infty$. The operating times of the components are limited, however, to the positive values. If a normal distribution is to be used for modelling the operating or down times of the components, then its truncated version can be used. The truncated normal distribution is given by

$$f(x) = \frac{1}{a\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}, \quad x \geq 0 \quad (2.62)$$

where a is the normalizing factor and is

$$\frac{1}{\sqrt{2\pi}} \int_{\frac{m}{\sigma}}^{\infty} e^{-u^2/2} du$$

The survivor function is given by

$$R(x) = \int_x^{\infty} f(u) du$$

and the hazard function

$$\phi(x) = \frac{f(x)}{R(x)}$$

The hazard rate of the normal distribution is a monotonically increasing function and the graphs of the probability density function, probability

distribution function and the hazard rate are shown in Fig. 2.2

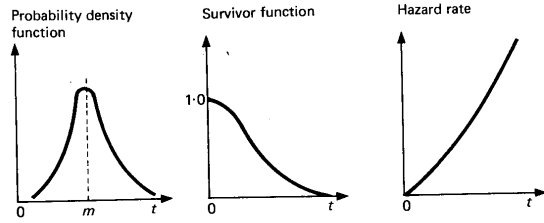


Fig. 2.2 Characteristics of the normal distribution

3 Log-normal Distribution

A non-negative random variable X is said to have a log-normal distribution with parameters m and σ if $Y = \log X$ is normally distributed with parameters m and σ . Since Y is a non-decreasing function of X

$$\begin{aligned} P(X \leq x) &= P(Y \leq y) \\ \text{or} \quad F_X(x) &= F_Y(y) \end{aligned}$$

Differentiating both sides

$$f_X(x) = f_Y(y) \frac{dy}{dx}$$

Therefore if

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-m)^2/2\sigma^2}$$

Then

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\log x - m)^2/2\sigma^2}, \quad x \geq 0 \quad (2.63)$$

which is the probability density function of X having a log-normal distribution. The corresponding cumulative distribution function is

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x e^{-(\log u - m)^2/2\sigma^2} \frac{du}{u}$$

Making the substitution

$$z = \frac{\log u - m}{\sigma}$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\log x - m} e^{-z^2/2} dz$$

which is the standard normal probability integral. The hazard rate is given by

$$\phi(x) = \frac{f(x)}{1 - F(x)}$$

The hazard rate of a log-normal distribution can be shown to increase to a maximum value and then decrease to zero as $x \rightarrow 0$. The log-normal distribution does not seem to be physically suited to model component lifetimes. It does however, seem to provide a reasonable fit for many component repair times.

The expression for the k th initial moment is easily derived as

$$m_k(X) = E(X^k) = E(e^{kY}) = e^{mk + \sigma^2 k^2/2} \quad (2.64)$$

The mean is

$$m_X = e^{m + \sigma^2/2} \quad (2.65)$$

and the variance

$$\sigma_X^2 = e^{2m + 2\sigma^2} - e^{2m + \sigma^2} \quad (2.66)$$

The graphs of the probability density function, probability distribution function and hazard rate are given in Fig. 2.3

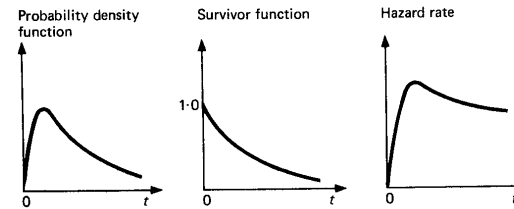


Fig. 2.3 Characteristics of the log-normal distribution

4 Weibull Distribution

The Weibull distribution is defined by the following probability density function

$$f(x) = \alpha \rho (\rho x)^{\alpha-1} e^{-(\rho x)^\alpha}, \quad x \geq 0 \quad (2.67)$$

The survivor function

$$\begin{aligned} R(x) &= \int_x^\infty \alpha \rho (\rho u)^{\alpha-1} e^{-(\rho u)^\alpha} du \\ &= e^{-(\rho x)^\alpha} \end{aligned} \quad (2.68)$$

and therefore the hazard function

$$\phi(x) = \alpha \rho (\rho x)^{\alpha-1} \quad (2.69)$$

The hazard rate increases with x for $\alpha > 1$ and decreases for $\alpha < 1$. The graphs of the probability density, probability distribution and the hazard rate of a Weibull distribution are given in Fig. 2.4. The k th moment of a Weibull distribution can be calculated by

$$E(X^k) = \int_0^\infty x^k \alpha \rho (\rho x)^{\alpha-1} e^{-(\rho x)^\alpha} dx$$

Substituting $z = (\rho x)^\alpha$

$$\begin{aligned} m_k &= \frac{1}{\rho^k} \int_0^\infty z^{k/\alpha} e^{-z} dz \\ &= \frac{1}{\rho^k} \Gamma\left(\frac{k}{\alpha} + 1\right) \end{aligned} \quad (2.70)$$

where $\Gamma(a)$ is a gamma function defined by

$$\Gamma(a) = \int_0^\infty z^{a-1} e^{-z} dz$$

When a is a positive integer

$$\Gamma(a) = (a-1)!$$

The mean value

$$m = \frac{\Gamma(1 + 1/\alpha)}{\rho} \quad (2.71)$$

The second initial moment

$$= \frac{\Gamma(1 + 2/\alpha)}{\rho^2}$$

and therefore variance

$$= \frac{\Gamma(1 + 2/\alpha) - (\Gamma(1 + 1/\alpha))^2}{\rho^2} \quad (2.72)$$

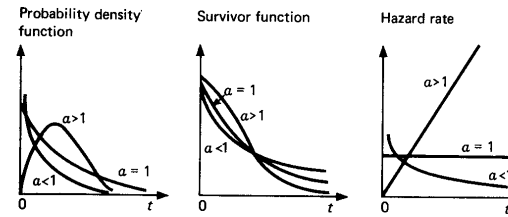


Fig. 2.4 Characteristics of the Weibull distribution

The Weibull distribution is often used in connection with mechanical components and has been used to model fatigue failure and the failure of ball bearings. It has also been recently used to describe the repair time of electric power generating units.

5 Gamma distribution

A non-negative continuous random variable X is said to be gamma distributed if its probability density function is given by

$$f(x) = \frac{\rho(\rho x)^{a-1} e^{-\rho x}}{\Gamma(a)} \quad (2.73)$$

This density is a function of parameters ρ and a both of which are positive constants. When a is an integer equal to a , then this distribution is also called the special Erlangian distribution,

$$f(x) = \frac{\rho(\rho x)^{a-1} e^{-\rho x}}{(a-1)!}$$

The probability distribution function is given by

$$F(x) = \int_0^x \frac{\rho(\rho u)^{a-1} e^{-\rho u}}{\Gamma(a)} du$$

Putting $\rho u = z$, $du = \frac{1}{\rho} dz$

$$F(x) = \frac{1}{\Gamma(a)} \int_0^{\rho x} z^{a-1} e^{-z} dz$$

The function $\int_0^{\rho x} z^{a-1} e^{-z} dz$ is called the incomplete gamma function. When a is an integer a (special Erlangian distribution), it can be shown by integrating by parts that

$$F(x) = 1 - \sum_{k=0}^{a-1} \frac{(\rho x)^k e^{-\rho x}}{k!}$$

The Laplace transform of a gamma function is given by

$$\bar{f}(s) = \left(\frac{\rho}{\rho + s} \right)^a$$

Therefore the mean $= (-1) \bar{f}'(s)|_{s=0}$

$$= \frac{a}{\rho}$$

The second initial moment $= (-1)^2 \bar{f}''(s)|_{s=0}$

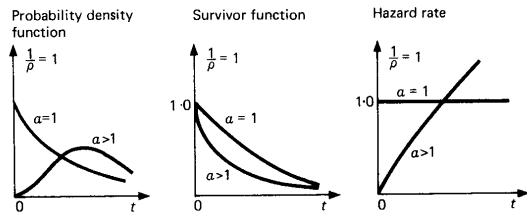


Fig. 2.5 Characteristics of the gamma distribution

$$= \frac{\alpha^2}{\rho^2} + \frac{\alpha}{\rho^2}$$

and the variance

$$= E(X^2) - (E(X))^2$$

$$= \frac{\alpha}{\rho^2}$$

Consequently the standard deviation is $\sqrt{\alpha}/\rho$ and the coefficient of variation is $1/\sqrt{\alpha}$. Some typical cases of the gamma function are shown in Fig. 2.5. The hazard rate is non-decreasing for $\alpha > 1$ and is bounded by ρ . For $0 < \alpha < 1$, the hazard rate is non-increasing and approaches zero as $x \rightarrow \infty$.

The gamma distribution is very useful due to its simplicity and also because it can be used to approximate many empirical distributions. A more detailed treatment of its usefulness is given in Chapter 6 while discussing the device of stages. A special case is when in Equation (2.73) $\rho = \frac{1}{2}$ and $\alpha = \frac{n}{2}$ where n is an integer. This distribution is the χ^2 (chi squared) distribution with n degrees of freedom.

Stochastic Processes

Consider the example of two transmission links again. $Z(t)$ defines the state of this system at time t . There is a random variable associated with each value of t . The family of random variables $(Z(t), t \geq 0)$ is called a stochastic process. The values assumed by the process are called the states of the system and the set of all possible states is called the state space. The set of the possible values of the indexing parameter is called the parameter space. The indexing parameter in the above example is time but other kinds of indexing parameter such as space are also possible. For example the number of fibres at a point on a yarn can be considered a stochastic process with the length of the yarn as the parameter. This book is, however, generally concerned with time dependent processes and therefore time is the basic indexing parameter. The stochastic process is, therefore, considered as a model of a system which develops in time according to probabilistic laws.

Consider two identical and independent systems of two transmission links. The state space for each system is defined as follows:

$$\begin{aligned} Z(t) = 0 & \quad \text{both links up} \\ Z(t) = 1 & \quad \text{one link up} \\ Z(t) = 2 & \quad \text{no link up} \end{aligned}$$

Assume that at time zero both the links are up, and the systems are observed

for twelve hours each. The state of each system as a function of time is shown in Fig. 2.6. These two observations are called the two independent realizations of the stochastic process. Realizations or sample paths could be obtained either by observing identical and independent systems or they could be constructed by appropriately using a table of random numbers or some equivalent randomizing device. The realizations could differ from one another in detail and in fact these differences are typical of random phenomenon. The idea of realizations often helps to provide a better insight in certain reliability problems.

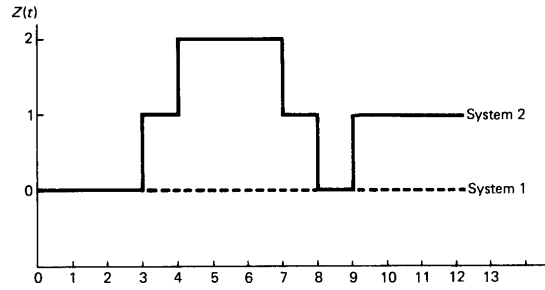


Fig. 2.6 Realizations of a stochastic process

The stochastic processes may be classified on the basis of the nature of the state space and the parameter space. The following are the four possible combinations:

1 Discrete state and parameter spaces

An example of this nature would be the number of successful missile flights in a missile firing scheme. The indexing parameter would be the number of missiles fired.

2 Discrete state space and continuous parameter space

Most of this book is concerned with stochastic processes of this kind. The states of a system of components with time as the indexing parameter is an example of this type of stochastic process.

3 Continuous state space and discrete parameter space

An example of this is the load on the electric power system observed every hour. The solution as such requires more sophisticated tools. These problems can, however, be often idealized by Category 1.

4 Continuous state space and continuous parameter space

An example of this nature is the storage in a dam observed as a function of time. Such problems can be idealized by Categories 1 or 2.

Probability Distributions

A stochastic process is defined for a set of points which may be either integer coordinates ($n = 0, 1, 2, \dots$) or an interval of real time ($0 \leq t$ or $-\infty < t < \infty$). At a particular point, the stochastic process is a simple random variable. Assuming k arbitrary time points $t_1, t_m, t_n, \dots, t_1 < t_m < t_n < \dots$, there are k random variables $Z(t_1), Z(t_m), Z(t_n), \dots$. In the discrete time case these may be denoted by $Z_1, Z_m, Z_n, \dots$. The stochastic process is completely determined in principle if the joint distribution of $Z_1, Z_m, Z_n, \dots$ is known for every k and every choice of l, m and n . In practice, however, it is rarely possible to work with these joint distributions and most of the information of interest can be obtained from transition distribution functions. One property of basic interest in the reliability evaluation of a system is the probability distribution of Z_n for the discrete time case and $Z(t)$ in the continuous time case.

Consider first the discrete time case. The stochastic process is said to be independent if

$$P(Z_n = x | Z_m = y, Z_l = z, \dots) = P(Z_n = x)$$

This means that the probability distribution of Z_n is independent of the present and the past history of the process. A slight weakening of this condition leads to the well known class of stochastic processes called the Markov process. In this case

$$P(Z_n = x | Z_m = y, Z_l = z, \dots) = P(Z_n = x | Z_m = y)$$

That is, the probability distribution of Z_n depends on the latest of the time points and none prior to that. For this reason, the Markov process is sometimes called memoryless. The Markov property essentially states that once the state occupied at a time point is known, the previous history of the process is not involved in determining the subsequent probability distributions. The Chapman-Kolmogorov equation gives the conditional probability density function for this process

$$P(Z_n = x | Z_l = z) = \int_{-\infty}^{\infty} P(Z_m = y | Z_l = z) P(Z_n = x | Z_m = y) dy \quad (2.74)$$

Equation (2.74) is for the continuous state space, discrete time case. The corresponding equation for continuous state space and continuous time can be written as

$$P(Z(t_n) \leq x | Z(t_l) = z) = \int_{-\infty}^{\infty} P(Z(t_n) \leq x | Z(t_m) = y) dP(Z(t_m) \leq y | Z(t_l) = z) \quad (2.75)$$

Equations (2.74) and (2.75) in their general form are rarely used in practice. They, however, convey the fundamental idea of recursively building the conditional probability density function over the long time interval (l, n) from those over the shorter time intervals (l, m) and (m, n) . If the conditional probability density function depends only on the distance $t_m - t_l$ and not on t_n and t_l the stochastic process is called time homogenous.

The remainder of this chapter discusses the 'discrete state space and discrete time' and 'discrete state space and continuous time' Markov processes. Most of the reliability modelling falls into the latter case. It is, however, sometimes convenient to idealize the continuous time by discrete time processes. The next chapter discusses the stochastic processes from the point of view of the frequency balance.

Markov Chains

This section considers a Markov process with discrete state space and discrete parameter space. Equation (2.74) can be simplified when the state space is discrete

$$P(Z_n = x | Z_l = z) = \sum_y P(Z_n = x | Z_m = y) P(Z_m = y | Z_l = z)$$

where x, y, z now denote the discrete states of the system. This equation develops the conditional probability density function over the longer interval from those of shorter interval. In practice, however, it is usual to work with one step transition probabilities. In this case the Markov property states

$$P(Z_n = x | Z_{n-1} = y, Z_{n-2} = z, \dots) = P(Z_n = x | Z_{n-1} = y) \quad (2.76)$$

If this one step transition probability is independent of n , i.e.

$$P(Z_n = x | Z_{n-1} = y) = P(Z_m = x | Z_{m-1} = y)$$

the process is time homogenous and transition probabilities are termed

stationary. The one step stationary probabilities will be denoted by p_{ij} , which is the probability of transiting from state i to state j in one step. It is easy to see that

$$\sum_j p_{ij} = 1$$

The n step transition probability $p_{ij}^{(n)}$ can be similarly defined as

$$p_{ij}^{(n)} = P(Z_{m+n} = j | Z_m = i)$$

The one step transition probabilities can be arranged in matrix form

$$P = (p_{ij})$$

This matrix is called the transition matrix and its ij th entry is the probability of transiting from state i to state j in one step. Each row sums to unity. A matrix which has non-negative entries with each row summing to 1 is called a stochastic matrix. Equation (2.75) can be written in terms of the single step transition probabilities as follows

$$p_{ij}^{(2)} = \sum_k p_{ik} p_{kj} \quad (2.77)$$

Arranged in matrix form

$$P^{(2)} = P \cdot P = P^2 \quad (2.78)$$

The matrix of two step transition probabilities can be found by squaring the transition matrix. It can be easily seen that

$$P^{(n)} = P^n \quad (2.79)$$

In practice, interest is usually focussed on the probability distribution of Z_n given the initial state of the system. The initial state of the system is defined by an initial probability vector

$$p^{(0)} = (p_0, p_1, p_2, \dots)$$

The vector of state probabilities after n steps is found by

$$p^{(n)} = p^{(0)} P^n \quad (2.80)$$

Written in the component form, this formula becomes

$$p_j^{(n)} = \sum_k p_k p_{kj}^{(n)} \quad (2.81)$$

where

$p_j^{(n)}$ = The probability of being in state j after n steps.

p_k = The probability of being in state k at the start.

and $p_{kj}^{(n)}$ = The probability of being in state j in n steps starting in state k .

Example: A person is practising firing. If he misses, he becomes nervous and the probability of the next shot being a hit reduces to $\frac{1}{2}$, but a hit bolsters his confidence and the chance of the next shot being a hit increases to $\frac{3}{4}$. If the initial shot is a hit, what is the probability of a hit on the fourth shot? Also calculate this probability for the initial shot being a miss.

Designating the hit by 0 and miss by 1, the object is to find the probability distribution for Z_4 . The state transition diagram is shown in Fig. 2.7.

From Equation (2.80)

$$p^{(3)} = p^{(0)} p^3$$

For the first shot being a hit

$$\begin{aligned} p^{(3)} &= (1 \quad 0) \begin{bmatrix} \frac{43}{64} & \frac{21}{64} \\ \frac{21}{32} & \frac{11}{32} \end{bmatrix} \\ &= \left(\frac{43}{64} \quad \frac{21}{64} \right) \end{aligned}$$

That is the probability of a hit on the fourth shot, the first shot being a hit, is $43/64$. If, however, the initial shot is a miss

$$\begin{aligned} p^{(3)} &= (0 \quad 1) \begin{bmatrix} \frac{43}{64} & \frac{21}{64} \\ \frac{21}{32} & \frac{11}{32} \end{bmatrix} \\ &= \left(\frac{21}{32} \quad \frac{11}{32} \right) \end{aligned}$$

The probability of a hit in this case is slightly less than the previous case.

It should be noted here that the (ij) th entry of $P^{(n)}$ represents the probability of being in the j th state after n steps, given the system started in state i . The states of a discrete Markov chain can be classified into the following types.

If the states i and j can be reached from each other in a finite number of steps, they are said to communicate. The set of states in which each pair of states communicate and which once entered cannot be left is called a closed communicating class. This is also called an ergodic set of states. On the other hand, a set of states in which every state can be reached from every other state is called a transient set. The discrete chain in which every state can be reached from every other state is termed irreducible or ergodic. In other words the states form a single closed communicating set. An ergodic chain in which each state can be entered only at certain periodic intervals is called cyclic or periodic chain. If a state exhibits this characteristic, then the state is termed periodic or cyclic. A discrete Markov chain which is ergodic and a-periodic is called a regular chain. The periodic chains and states are troublesome to deal with, but fortunately reliability problems are most frequently described by regular chains.

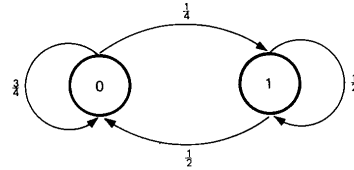


Fig. 2.7 State transition diagram

Equilibrium Distribution

In the firing practice example

$$P^3 = \begin{bmatrix} 0.6719 & 0.3281 \\ 0.6563 & 0.3438 \end{bmatrix}$$

$$P^6 = \begin{bmatrix} 0.6667 & 0.3333 \\ 0.6665 & 0.3335 \end{bmatrix}$$

$$P^{12} = \begin{bmatrix} 0.6666 & 0.3334 \\ 0.6666 & 0.3334 \end{bmatrix}$$

It can be seen that the entries of P^n seem to be approaching a limiting value. Is this true in all cases? The following results are stated without proof.

1. In any Markov chain which is not cyclic the limit $x_j = \lim_{n \rightarrow \infty} p_j^{(n)}$ exists.
2. In any a-periodic, irreducible Markov chain the above limit does not depend on the initial probability distribution so that

$$x_j = \lim_{n \rightarrow \infty} p_j^{(n)} = \lim_{n \rightarrow \infty} p_{i_j}^{(n)}$$

3. In a finite regular Markov chain, each row approaches a stationary probability vector $\alpha = (\alpha_0, \alpha_1, \dots)$. This is called the unique stationary probability vector of the process and

$$\alpha P = \alpha \quad (2.82)$$

This relationship is very useful for determining the limiting state probabilities (also called steady state probabilities) of the process. In the firing practice example

$$(\alpha_0 \quad \alpha_1) \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = (\alpha_0 \quad \alpha_1)$$

i.e.

$$-\frac{1}{4}\alpha_0 + \frac{1}{2}\alpha_1 = 0 \quad (2.83)$$

$$\frac{1}{4}\alpha_0 - \frac{1}{2}\alpha_1 = 0 \quad (2.84)$$

These two equations are identical, therefore an equation of the following form can be used

$$\alpha_0 + \alpha_1 = 1 \quad (2.85)$$

From Equations (2.83) and (2.85)

$$\alpha_0 = \frac{2}{3}$$

and

$$\alpha_1 = \frac{1}{3}$$

It can be seen that these values could also be obtained by multiplying P , a large number of times.

Time Specific Behaviour

It has been shown that the n -step probability distribution of the discrete Markov chain can be found from P^n where P is the transition matrix. In determining higher powers of P , the following matrix product is often useful

$$P^n = P^{n-m} P^m \quad (2.86)$$

The multiplication of large matrices is quite unwieldy using hand calculations but easily accomplished when a digital computer is used. Though this method of matrix multiplication is quite useful, the following technique can be used for very large powers of P .

If the matrix P has N distinct real eigenvalues, then it can be proved that there exists a matrix S having an inverse S^{-1} such that

$$SPS^{-1} = D \quad (2.87)$$

The matrix P is then said to be similar to the diagonal matrix D . In the diagonal matrix all but the diagonal elements are zero. The diagonal elements of D are the eigenvalues of P and can be determined from the following relationship.

$$\det(P - dI) = 0 \quad (2.88)$$

Equation (2.87) can be rearranged as

$$SP = DS \quad (2.89)$$

The row s_i of S is termed the i th left eigenvector associated with the eigenvalue d_{ii} . Similarly rearranging (2.89)

$$PS^{-1} = S^{-1}D \quad (2.90)$$

The j th column vector of S^{-1} is termed the j th right eigenvector of S . It can be seen from (2.87) that

$$SPS^{-1}SPS^{-1} = SP^2S^{-1} = D^2$$

and by induction

$$P^n = S^{-1}D^nS \quad (2.91)$$

The n th power of P can therefore be found from the n th power of the diagonal

matrix which is easy. The difficult part, however, is to determine the eigenvalues and the associated eigenvectors of P . Several numerical techniques are available for determining these elements. In many practical problems, the basic matrix multiplication technique is quite adequate. The matrix algebra approach to the firing practice problem is as follows

$$P = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Therefore

$$(P - dI) = \begin{bmatrix} \frac{3}{4} - d & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} - d \end{bmatrix}$$

The eigenvalues can now be found by equating the determinant to zero.

$$\begin{vmatrix} \frac{3}{4} - d & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} - d \end{vmatrix} = 0$$

$$(\frac{3}{4} - d)(\frac{1}{2} - d) - \frac{1}{8} = 0$$

i.e.

$$4d^2 - 5d + 1 = 0$$

The roots of this equation give the eigenvalues of P

$$d_0 = 1$$

$$d_1 = \frac{1}{4}$$

Therefore

$$D = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$$

The next step involves determining the left and right eigenvectors of P . From Equation (2.89)

$$\begin{aligned} \text{i.e. } \frac{3}{4}s_{00} + \frac{1}{4}s_{01} &= s_{00} \\ -\frac{1}{4}s_{00} + \frac{1}{4}s_{01} &= 0 \end{aligned} \quad (2.93)$$

Also

$$\frac{1}{4}s_{00} - \frac{1}{2}s_{01} = 0 \quad (2.94)$$

$$\frac{1}{2}s_{10} + \frac{1}{2}s_{11} = 0 \quad (2.95)$$

and

$$\frac{1}{4}s_{10} + \frac{1}{4}s_{11} = 0 \quad (2.96)$$

Equation (2.93) and (2.94) are identical, as are Equations (2.95) and (2.96). There are now two equations and four unknowns and therefore the magnitudes of the eigenvectors cannot be uniquely determined. Assuming

$$s_{00} = s_{11} = 1$$

$$s_{01} = \frac{1}{2}$$

$$s_{10} = -1$$

Therefore

$$S = \begin{bmatrix} 1 & \frac{1}{2} \\ -1 & 1 \end{bmatrix}$$

It should be noted that the elements of S , in general, are not determined uniquely. Each row of S is determined up to a multiplicative constant. S^{-1} can be found by inverting S

$$S^{-1} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Using (2.91)

$$\begin{aligned} P^n &= \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} (1)^n & 0 \\ 0 & (\frac{1}{4})^n \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{2}{3} + \frac{1}{3}(\frac{1}{4})^n & \frac{1}{3} - \frac{1}{3}(\frac{1}{4})^n \\ \frac{2}{3} - \frac{2}{3}(\frac{1}{4})^n & \frac{1}{3} + \frac{2}{3}(\frac{1}{4})^n \end{bmatrix} \end{aligned}$$

For $n = 3$

$$P^3 = \begin{bmatrix} \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{64} & \frac{1}{3} - \frac{1}{3} \cdot \frac{1}{64} \\ \frac{2}{3} - \frac{2}{3} \cdot \frac{1}{64} & \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{64} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{43}{64} & \frac{21}{64} \\ \frac{21}{32} & \frac{11}{32} \end{bmatrix}$$

as previously found by the matrix multiplication approach. Similarly

$$P^n_{n \rightarrow \infty} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

First Passage Times

One parameter of interest in many Markov Chain problems is the time to encounter a state for the first time. This is called the first passage time. If this state is an absorbing state or has been made an absorbing state, this is called the time of absorption. In reliability engineering this concept is used to calculate the mean time to first failure, MTTF. As noted earlier, almost all the cases of practical interest are regular chains, i.e. chains in which all the states communicate and which are not cyclic. In these cases, the mean first passage times and their variance can be obtained from the fundamental matrix Z defined as below

$$Z = [I - [P - A]]^{-1} \quad (2.97)$$

where

I is the identity matrix

P is the transition matrix

and A is the matrix each row of which is the limiting probability vector

$$\alpha = (\alpha_0, \alpha_1, \dots)$$

The mean first passage time matrix $\bar{T}$ is given by

$$\bar{T} = [I - Z + UZ_d]D \quad (2.98)$$

where

$\bar{T}$ is the mean first passage time matrix such that $\bar{t}_{ij}$ represents the mean time or mean number of steps to go from state i to j

U A unit matrix, i.e. with all entries 1

Z_d Matrix resulting from Z by setting off-diagonal elements equal to zero

D Diagonal matrix such that $d_{ii} = 1/\alpha_i$

The variance of the first passage times can also be explicitly determined. Denoting the first passage time from state i to j by t_{ij} , define the matrix W

$$W = (E(t_{ij}^2))$$

This matrix can be computed from the fundamental matrix by

$$W = \bar{T}(2Z_d D - I) + 2[Z\bar{T} - U(Z\bar{T})_d] \quad (2.99)$$

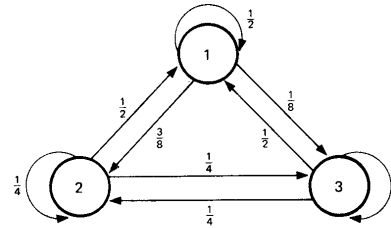
$(Z\bar{T})_d$ is a matrix obtained by setting the off-diagonal elements of $Z\bar{T}$ equal to zero.

The variance can now be obtained using Equation (2.13)

$$V(t_{ij}) = E(t_{ij}^2) - (E(t_{ij}))^2$$

$$= w_{ij} - \bar{t}_{ij}^2$$

Example: A discrete Markov chain has the state transition diagram shown below. Find the matrix of the mean first passage times.



The transition matrix for this state space diagram is

$$P = \begin{bmatrix} \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

The vector α can be computed by

$$(\alpha_1 \quad \alpha_2 \quad \alpha_3) \begin{bmatrix} \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} = (\alpha_1 \quad \alpha_2 \quad \alpha_3)$$

On solving these equations along with $\alpha_1 + \alpha_2 + \alpha_3 = 1$

$$\alpha_1 = \frac{1}{2}$$

$$\alpha_2 = \frac{5}{16}$$

$$\alpha_3 = \frac{3}{16}$$

Therefore

$$A = \begin{bmatrix} \frac{1}{2} & \frac{5}{16} & \frac{3}{16} \\ \frac{1}{2} & \frac{5}{16} & \frac{3}{16} \\ \frac{1}{2} & \frac{5}{16} & \frac{3}{16} \end{bmatrix}$$

$$I - [P - A] = \begin{bmatrix} 1 & -\frac{1}{16} & \frac{1}{16} \\ 0 & \frac{17}{16} & -\frac{1}{16} \\ 0 & \frac{1}{16} & \frac{15}{16} \end{bmatrix}$$

Z can be now determined from the inverse of the above matrix

$$Z = \begin{bmatrix} 1 & \frac{1}{16} & -\frac{1}{16} \\ 0 & \frac{15}{16} & \frac{1}{16} \\ 0 & -\frac{1}{16} & \frac{17}{16} \end{bmatrix}$$

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{16}{5} & 0 \\ 0 & 0 & \frac{16}{3} \end{bmatrix}$$

Substituting into (2.98)

$$\bar{T} = \begin{bmatrix} 1 & \frac{14}{16} & \frac{18}{16} \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{16}{5} & 0 \\ 0 & 0 & \frac{16}{3} \end{bmatrix}$$

$$= \begin{bmatrix} 2 & \frac{14}{5} & 6 \\ 2 & \frac{16}{5} & \frac{16}{3} \\ 2 & \frac{16}{5} & \frac{16}{3} \end{bmatrix}$$

Alternative Approach to First Passage Times

The technique for calculating the mean and variance of first passage times for regular Markov chains has been illustrated. It is also possible to calculate these quantities by making state j an absorbing state and applying the theory of absorbing chains. An absorbing chain is one which once entered cannot be left. The behaviour of the stochastic process before once hitting state j will be the same as that of the original process. The first passage time from state i to state j is now the time of absorption starting from state i in the new process. The basic results for this absorbing chain can be obtained from the fundamental matrix N

$$N = [I - Q]^{-1} \quad (2.100)$$

where

N = The fundamental matrix whose n_{ik} denotes the mean number of times the process is in state k before absorption, the process having been started in state i .

Q = The matrix obtained by deleting the j th row and the j th column from matrix P of transition probabilities.

The mean first passage time from state i to j is therefore

$$\bar{t}_i = \sum_{k=1}^{N-1} n_{ik}$$

The variance column vector is given by

$$W = [2N - I] \bar{t} - \bar{t}_s \quad (2.101)$$

where

- W_i = The variance of the first passage time from state i to state j (the one made into an absorbing state).
 $\bar{t}$ = The column vector such that $\bar{t}_i$ is the mean first passage time from i to j .
 $\bar{t}_s$ = The column vector with $\bar{t}_{si} = \bar{t}_i^2$

It can be seen that this approach gives additional information about the mean first passage time by providing the components spent in various states before once hitting state j . This method can be illustrated by application to the previous example. Determine the mean first passage times from states 1 and 2 to state 3. Truncating the third column and row

$$I - Q = \begin{bmatrix} \frac{1}{2} & -\frac{3}{8} \\ -\frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

The fundamental matrix N can be found from its inverse

$$N = \begin{bmatrix} 4 & 2 \\ \frac{8}{3} & \frac{8}{3} \end{bmatrix}$$

Starting in state 1, the process visits states 1 and 2, 4 and 2 times before first hitting state 3. Therefore

$$\bar{t}_{13} = 4 + 2 = 6$$

and

$$\bar{t}_{23} = \frac{8}{3} + \frac{8}{3} = 5.333$$

It can be seen that these entries agree with the elements of $\bar{T}$ found previously. If state 3 was considered to be the failed state of the system, then MTTF is 6 when state 1 is taken as the initial state.

Continuous Parameter Markov Chains

Many of the problems encountered in system reliability can be modelled using continuous parameter Markov chains. The next chapter examines frequency balancing techniques as an alternative way of looking at the stochastic process.

For $u < v < t$, the Markov property for a continuous parameter Markov chain would be

$$P(Z(t) = k | Z(v) = j, Z(u) = i) = P(Z(t) = k | Z(v) = j)$$

This property is basically of the form

$$P(Z(t+x) = j | Z(t) = i)$$

and is termed as the probability of transition from state i to state j during the time interval t to $t+x$. If this transition probability does not depend on the initial time t but only on the elapsed time x , then the process is said to be time homogeneous. This book is primarily concerned with this class of process. The transition probability will be denoted by

$$p_{ij}(x) = P(Z(t+x) = j | Z(t) = i)$$

for any x . The Chapman-Kolmogorov Equation (2.75) can now be written as

$$p_{ij}(t+x) = \sum_k p_{ik}(t) p_{kj}(x) \quad (2.102)$$

The transition probabilities must satisfy the following conditions

$$0 \leq p_{ij}(x) \leq 1 \quad (2.103)$$

and

$$\sum_j p_{ij}(x) \leq 1 \quad (2.104)$$

In Equation (2.104), if $\sum_j p_{ij}(x) = 1$ for all i and x , then the process is

called honest but if the inequality holds then there is non-zero probability of the process escaping to infinity and such a process is termed dishonest. This book is concerned only with honest processes. In the case of discrete parameter chains, the basic elements are the one step transition probabilities. In a continuous parameter case the equivalent elements are the limiting values, i.e. as $x \rightarrow 0$. Define the transition intensity or rate as

$$\begin{aligned} \lambda_{ij} &= \left. \frac{dp_{ij}(x)}{dx} \right|_{x=0} \\ & \quad i \neq j \\ &= \lim_{\Delta x \rightarrow 0} \frac{p_{ij}(\Delta x) - 0}{\Delta x} \end{aligned}$$

$$\text{i.e. } p_{ij}(\Delta x) = \lambda_{ij}\Delta x + 0(\Delta x) \quad (2.105)$$

for $i = j$

$$\begin{aligned} \lambda_{ij} &= \left. \frac{dp_{ij}(x)}{dx} \right|_{x=0} \\ &= \lim_{\Delta x \rightarrow 0} \frac{p_{ij}(\Delta x) - 1}{\Delta x} \end{aligned}$$

$$\text{i.e. } p_{ii}(\Delta x) = \lambda_{ii}\Delta x + 1 + 0(\Delta x) \quad (2.106)$$

Differentiating both sides of (2.104) for equality and setting $x = 0$

$$\begin{aligned} \lambda_{ii} + \sum_{j \neq i} \lambda_{ij} &= 0 \\ \text{i.e. } \lambda_{ii} &= - \sum_{j \neq i} \lambda_{ij} \end{aligned}$$

Therefore

$$p_{ii}(\Delta x) = 1 - \sum_{j \neq i} \lambda_{ij}\Delta x + 0(\Delta x) \quad (2.107)$$

In Equation (2.105), $p_{ij}(\Delta x)$ represents the probability of transiting from state i to state j during the interval of length Δx and this is equal to $\lambda_{ij}\Delta x$ plus a term which when divided by Δx tends to zero as $\Delta x \rightarrow 0$. Equation (2.107) can be interpreted in a similar manner. Equation (2.102) can now be written for a small increment of time Δt as

$$\begin{aligned} p_{ij}(t + \Delta t) &= \sum_k p_{ik}(t) p_{kj}(\Delta t) \\ &= p_{ij}(t) p_{jj}(\Delta t) + \sum_{k \neq j} p_{ik}(t) p_{kj}(\Delta t) \end{aligned} \quad (2.108)$$

where

$$p_{ij}(t) = P(Z(t) = j | Z(0) = i)$$

Substituting from (2.105) and (2.106)

$$p_{ij}(t + \Delta t) = p_{ij}(t)(1 + \lambda_{jj}\Delta t) + \sum_{k \neq j} p_{ik}(t)\lambda_{kj}\Delta t + 0(\Delta t)$$

$$\text{i.e. } \frac{p_{ij}(t + \Delta t) - p_{ij}(t)}{\Delta t} = p_{ij}(t)\lambda_{jj} + \sum_{k \neq j} p_{ik}(t)\lambda_{kj} + \frac{0(\Delta t)}{\Delta t}$$

and as $\Delta t \rightarrow 0$

$$p'_{ij}(t) = \sum_k p_{ik}(t)\lambda_{kj}$$

If $P'_i(t)$ denotes the row vector whose j th element is $p'_{ij}(t)$, i.e. the probability of being in the j th state at time t given that the process was initially in state i , then the above equation can be written as

$$P'_i(t) = P_i(t)R \quad (2.109)$$

where R is the transition rate matrix whose ij th element is λ_{ij} . In a more general form Equation (2.109) becomes

$$P'(t) = P(t)R \quad (2.110)$$

where $P(t)$ has $p_{ij}(t)$ as its (ij) th element. The initial condition for (2.110) is

$$P(0) = I$$

If, however the initial state of the system is defined by a probability distribution in the form of a row vector $p(0)$, the distribution at t is given by $p(0)P$. The system of equations (2.110) is termed as the system of forward equations.

At this point it is interesting to probe a little into the significance of the transition rates. Let X_{kj} be a random variable defining the duration of state k under the condition that the next transition will be to state j . In accordance with Equation (2.7), the hazard rate is

$$\phi_{kj}(x) = \lim_{\Delta x \rightarrow 0} \frac{P[x < X_{kj} \leq x + \Delta x | x < X_{kj}]}{\Delta x}$$

i.e.

$$P[x < X_{kj} \leq x + \Delta x | x < X_{kj}] = \phi_{kj}(x)\Delta x + 0(\Delta x)$$

The left hand side can be interpreted as $p_{kj}(\Delta x)$ if the process has been in state k for time x . If the process is to be Markovian then $\phi_{kj}(x)$ must be

independent of x as the process is independent of the past. Therefore

$$p_{kj}(\Delta x) = \phi_{kj}\Delta x + o(\Delta x)$$

Comparing with (2.105).

$$\lambda_{kj} = \phi_{kj}$$

That is the transition rate λ_{kj} is the hazard rate of the random variable defining the duration of state k under the condition of transiting to state j . The exponential is the only distribution having a constant hazard rate and therefore the random variables underlying the time homogenous Markov process must be exponentially distributed.

Although time homogenous Markov Chains are the main interest in system reliability, there is no additional difficulty in extending the above arguments to transition rates which are functions of system time, i.e. when λ_{ij} is $\lambda_{ij}(t)$. The process, however, becomes non-Markovian when the transition rates are a function of the state residence times. These processes are treated in Chapter 6. In the case when the transition rates are functions of system time t there exists a family of matrices $P(u, t)$, for $t > u$ whose elements are

$$p_{ij}(u, t) = P(X(t) = j | X(u) = i)$$

In this case however the transition probability depends not only on $(t - u)$ but on u as well. The system of forward equations can now be written as

$$\frac{\partial P(u, t)}{\partial t} = P(u, t)R(t) \quad (2.111)$$

This equation is called the Kolmogorov differential equation.

Transient Behaviour

Equation (2.110) is a system of linear differential equations with constant coefficients. If the eigenvalues of R are distinct, the solution of Equation (2.110) can be easily obtained in the form

$$P(t) = SD(t)S^{-1} \quad (2.112)$$

where

- $D(t)$ = The diagonal matrix whose (ii) th element is $\exp\{r_i t\}$, r_i being the i th eigenvalue of R
- S = The matrix formed by right eigenvectors of R

S^{-1} = The matrix formed either by inverting S or from the left eigenvalue of R

The proof of Equation (2.112) may be found in books on differential equations. In practice if t is short, the solution may be found by the following technique.

$$P'(t) = P(t)R$$

As $\Delta t \rightarrow 0^+$

$$\begin{aligned} P(t + \Delta t) &= P(t) + P'(t)\Delta t \\ &= P(t)(I + R\Delta t) \end{aligned}$$

It time t is divided into a very large number of equal intervals Δt , so that Δt is very small (≈ 0), the above expression can be written as a recursive relationship

$$P(j\Delta t) = P(j-1\Delta t)[I + R\Delta t] \quad (2.113)$$

It should be noted that (2.113) implies the approximation of a Markov process in continuous time by a discrete time Markov process with steps equal to Δt . The (ij) th element of $[I + R\Delta t]$ is $\lambda_{ij}\Delta t$, i.e. the probability of transiting from state i to state j in one step of length Δt . Therefore $[I + R\Delta t]$ is a one step transition probability matrix. It can also be seen that

$$P(j\Delta t) = [I + R\Delta t]^j$$

which is the matrix multiplication technique in the discrete time case

Equilibrium Probability Distribution

As $t \rightarrow \infty$, the probability distribution of $Z(t)$ tends to an equilibrium distribution. For all processes having a finite number of mutually communicating states, the unique solution can be found by solving $(N-1)$ equations from

$$pR = 0 \quad (2.114)$$

and

$$\sum_i p_i = 1 \quad (2.115)$$

where p is a row vector whose i th element p_i is the steady state probability of being in the i th state. It can be proved that p_i equals the expected value of the proportion of a long realization spent in state i . In most cases the steady state

probabilities are the only quantities of interest.

Example: One of the processes commonly encountered in reliability studies is the two state Markov process. The state transition diagram for this process is shown below.

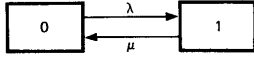


Fig. 2.8 Two-state Markov process

The transition rate matrix

$$R = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

Taking the Laplace of Equation (2.110)

$$sP(s) - P(0) = P(s)R$$

i.e.

$$P(s) = P(0)[sI - R]^{-1}$$

For the two state process

$$\begin{aligned} P(s) &= P(0) \begin{bmatrix} s + \lambda & -\lambda \\ -\mu & s + \mu \end{bmatrix}^{-1} \\ &= \frac{1}{s(s + \lambda + \mu)} \begin{bmatrix} s + \mu & \lambda \\ \mu & s + \lambda \end{bmatrix} \\ P(t) &= \begin{bmatrix} \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} & \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t} \\ \frac{\mu}{\lambda + \mu} - \frac{\mu}{\lambda + \mu} e^{-(\lambda + \mu)t} & \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-(\lambda + \mu)t} \end{bmatrix} \end{aligned}$$

The probability vector $p(t)$ of the state probabilities can be obtained by

$$(p_0(t) \ p_1(t)) = (p_0(0) \ p_1(0))P(t)$$

Therefore

$$\begin{aligned} p_0(t) &= \frac{\mu}{\lambda + \mu} (p_0(0) + p_1(0)) + \left(p_0(0) \frac{\lambda}{\lambda + \mu} - p_1(0) \frac{\mu}{\lambda + \mu} \right) e^{-(\lambda + \mu)t} \\ &= \frac{\mu}{\lambda + \mu} + \left(p_0(0) - \frac{\mu}{\lambda + \mu} \right) e^{-(\lambda + \mu)t} \end{aligned} \quad (2.116)$$

and

$$p_1(t) = \frac{\lambda}{\lambda + \mu} + \left(p_1(0) - \frac{\lambda}{\lambda + \mu} \right) e^{-(\lambda + \mu)t} \quad (2.117)$$

As $t \rightarrow \infty$

$$p_0(t) = p_0 = \frac{\mu}{\lambda + \mu}$$

and

$$p_1(t) = p_1 = \frac{\lambda}{\lambda + \mu}$$

The steady state solution can also be obtained by the application of Equations (2.114) and (2.115)

$$(p_0 \ p_1) \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} = 0$$

Therefore

$$-\lambda p_0 + \mu p_1 = 0$$

and

$$\lambda p_0 - \mu p_1 = 0$$

One of these identical equations can be used with

$$p_0 + p_1 = 1$$

to give

$$p_0 = \frac{\mu}{\lambda + \mu}$$

and

$$p_1 = \frac{\lambda}{\lambda + \mu}$$

The probabilities p_0 and p_1 are independent of the initial condition. This seems to be intuitively true because after a long time many transitions between 0 and 1 would have taken place and therefore the effect of the initial condition tends to diminish. The probabilities p_0 and p_1 can be interpreted in two ways. The first interpretation can be in terms of an average taken over a large number of the realizations taken at a single point in time. If out of n realizations, the process is n_0 times in state 0 at a time t remote from the time origin then

$$p_0 = \frac{n_0}{n} \quad n \rightarrow \infty$$

The second interpretation is in terms of the limiting proportion of time spent in state 0 in a single long realization. The parameters λ and μ are the hazard rates of exponential distributions and therefore they are the reciprocals of the mean time spent in state 0 and state 1, i.e.

$$\lambda = \frac{1}{E(X_0)}$$

$$\mu = \frac{1}{E(X_1)}$$

where X_0 and X_1 are the random variables denoting durations of 0 and 1 state.

Considering $2n$ transitions in a single long realization of the stochastic process, the process will be n times in state 0 and n times in state 1. Therefore, the proportion of time spent in state 0 in $2n$ transitions is

$$R_0^{(n)} = \frac{X_{01} + X_{02} + \dots + X_{0n}}{(X_{01} + X_{02} + \dots + X_{0n}) + (X_{11} + X_{12} + \dots + X_{1n})}$$

Dividing both numerator and denominator by n

$$R_0^{(n)} = \frac{\bar{X}_0}{\bar{X}_0 + \bar{X}_1}$$

As the number of transitions tends to be large, the average values tend to the means (law of large numbers) and therefore

$$R_0^{(n)} = \frac{E(X_0)}{E(X_0) + E(X_1)} \quad n \rightarrow \infty$$

$$\begin{aligned} &= \frac{\mu}{\lambda + \mu} \\ &= p_0 \end{aligned}$$

Therefore p_0 is the limiting proportion of the time spent in state 0 in a single long realization of the two state stochastic process. A similar interpretation holds for state 1.

First Passage Times

Denote the first passage time from state i to state j by T_{ij} , i.e. this is the time to enter state j for the first time starting in state i . If the state j is now made an absorbing state, the behaviour of the new stochastic process and the original process is the same until meeting j for the first time. If $p_{ij}(t)$ is the probability of being in state j , starting in state i for the new process then

$$P(T_{ij} \leq t) = p_{ij}(t)$$

The probability density function $f_{ij}(t)$ can be found by differentiation

$$f_{ij}(t) = \frac{d}{dt} (P(T_{ij} \leq t)) = \frac{d}{dt} p_{ij}(t)$$

The Laplace transform can be obtained by

$$\bar{f}_{ij}(s) = s\bar{p}_{ij}(s) \quad (2.118)$$

The bar indicates a Laplace transform. After evaluating the right hand side, the explicit density function can be obtained by inversion. The moments of the first passage times can be found by referring to Equation (2.43).

The k th moment of the first passage time can be found by differentiating Equation (2.43), k times

$$T_{ij}^{(k)} = (-1)^k \frac{d^k}{ds^k} \bar{f}_{ij}(s) \Big|_{s=0} \quad (2.119)$$

If the absorbing state is the failed state, then the mean first passage time represents the MTTFF. The above procedure can be conveniently carried out in the matrix form. Let the states 1 to J be the elements of subset X^+ and $J+1$ to

N be the elements of X^- . It is required to find the first passage time to the subset X^- . The matrix of transition rates can now be partitioned as follows

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

where

R_{11} is a $J \times J$ matrix

R_{12} is a $J \times (N - J)$ matrix

R_{21} is a $(N - J) \times J$ matrix

and

R_{22} is a $(N - J) \times (N - J)$ matrix

The states jeX^- are now absorbing states and therefore R_{21} and R_{22} are set to zero. Let $p(t)$ be the vector of state probabilities for an initial starting condition. This vector can be expressed as $(p_+(t) \ p_-(t))$ where $p_+(t)$ and $p_-(t)$ are the vectors containing the states ieX^+ and ieX^- respectively. The forward differential equation now becomes

$$\frac{d}{dt}(p_+(t) \ p_-(t)) = (p_+(t) \ p_-(t)) \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}$$

Therefore

$$p'_+(t) = p_+(t) R_{11}$$

and

$$p'_-(t) = p_+(t) R_{12}$$

Taking the Laplace transforms

$$s\bar{p}_+(s) - p_+(0) = \bar{p}_+(s) R_{11}$$

$$s\bar{p}_-(s) = \bar{p}_+(s) R_{12}$$

$p_-(0) = 0$ as the process started in $i \in X^+$. These equations can be rearranged as

$$\bar{p}_-(s) = p_+(0)[sI - R_{11}]^{-1} R_{12} \quad (2.120)$$

and

$$s\bar{p}_-(s) = p_+(0)(sI - R_{11})^{-1} R_{12} \quad (2.121)$$

The probability of being in subset X^- at time t is $p_-(t) U_{N-k}$ where U_{N-k} is a unit vector of dimension $N-k$. From Equation (2.118) it can be seen that the

Laplace transform of the probability density function of the first passage time is

$$\begin{aligned} \tilde{f}(s) &= s\bar{p}_-(s) U_{N-k} \\ &= p_+(0)(sI - R_{11})^{-1} R_{12} U_{N-k} \end{aligned}$$

Since the rows of the transition rate matrix sum to zero

$$R_{12} U_{N-k} = R_{11} U_k$$

Therefore

$$\tilde{f}(s) = p_+(0)[sI - R_{11}]^{-1} R_{11} U_k \quad (2.122)$$

The r th initial moment can be found by Equation (2.119)

$$T^{(h)} = k! p_+(0)(-R_{11})^{-h} U_k \quad (2.123)$$

The mean is

$$\bar{T} = T^{(1)} = p_+(0)(-R_{11})^{-1} U_k \quad (2.124)$$

If the process started in the first state

$$\bar{T} = (1 \ 0 \ 0 \ \dots \ 0)(-R_{11})^{-1} U_k$$

If X^- represents the failed condition $\bar{T}$, then $\bar{T}$ is the MTTF. It should be realized that Equation (2.124) can be derived from the theory of discrete time Markov chains by assuming that each step of the chain is $\Delta t \approx 0$. The matrix of one step transition probabilities becomes $[I + R \Delta t]$ and by truncating the absorbing states $Q = [I + R_{11} \Delta t]$ and therefore the fundamental matrix

$$N = [I - Q]^{-1} = \frac{1}{\Delta t} [-R_{11}]^{-1}$$

This matrix gives the number of steps spent in the different states. The time spent in the different states can be obtained by multiplying by Δt i.e. the step length. From this point on it is easy to see that

$$\bar{T} = p_+(0)(-R_{11})^{-1} U_k$$

The first passage time represents the time of entering a state or a set of states for the first time, starting in a particular state. It is sometimes necessary, however, to find the mean time spent in subset X^+ or X^- . For example, if X^+ and X^- represent the up and down states respectively, these time

parameters represent the mean up time and the mean down time. In order to calculate these quantities, it is necessary to know the probabilities of beginning X^+ in the various states, which are its elements. Denoting the steady state probabilities of being in various states of the original process by p_i , the probability of beginning X^+ in state j is

$$p_j(0) = \frac{\sum_{i \in X^-} p_i \lambda_{ij}}{\sum_{j \in X^+} \sum_{i \in X^-} p_i \lambda_{ij}}$$

In vector form

$$p_+(0) = \frac{p_- R_{21}}{p_- R_{21} U_k}$$

In the steady state

$$p_+ R_{11} + p_- R_{21} = 0$$

Therefore

$$p_+(0) = \frac{-p_+ R_{11}}{p_+ R_{11} U_k} = \frac{-p_+ R_{11}}{p_+ R_{12} U_{N-k}}$$

Substituting in (2.124), the mean stay in X^+ is

$$\begin{aligned} T^+ &= \frac{-p_+ R_{11} (-R_{11})^{-1} U_k}{p_+ R_{12} U_{N-k}} = \frac{p_+ U_k}{p_+ R_{12} U_{N-k}} \\ &= \sum_{i \in X^+} p_i / \sum_{i \in X^+} p_i \sum_{j \in X^-} \lambda_{ij} \\ &= \sum_{i \in X^+} p_i / \sum_{i \in X^-} p_i \sum_{j \in X^+} \lambda_{ij} \end{aligned} \quad (2.125)$$

In a similar manner the mean duration in state X^-

$$\begin{aligned} T^- &= \sum_{i \in X^-} p_i / \sum_{i \in X^-} p_i \sum_{j \in X^+} \lambda_{ij} \\ &= \sum_{i \in X^-} p_i / \sum_{i \in X^+} p_i \sum_{j \in X^-} \lambda_{ij} \end{aligned} \quad (2.126)$$

The mean cycle time, i.e. the time between two successive encounters of X^+ or X^-

$$\begin{aligned} T &= T^+ + T^- \\ &= 1 / \sum_{i \in X^-} p_i \sum_{j \in X^+} \lambda_{ij} \\ &= 1 / \sum_{i \in X^+} p_i \sum_{j \in X^-} \lambda_{ij} \end{aligned} \quad (2.127)$$

In the next chapter these relationships will be derived from the frequency balancing technique.

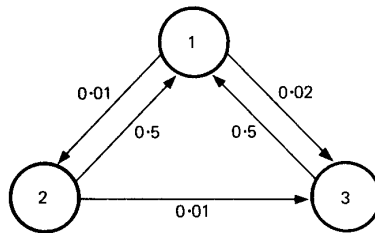
Exercises

1. X is a non-negative continuous random variable such that conditional on X being greater than a fixed value $t > 0$, the probability density function of $X - t$ is the same as the unconditional probability density function of X . Prove that X has a negative exponential probability density function.
2. Assume that X is normally distributed with $m = 0.4$ and $\sigma = 4$, find
 - (a) $P(X > 1.5)$
 - (b) $P(X \leq 0.5)$
 - (c) $P(-3 < X < 1)$
3. Suppose that $X_i, i = 1, 2, \dots, n$ are independent random variables, gamma distributed with parameters α_i, ρ . Prove that the random variable $\sum_{i=1}^n X_i$ is also gamma distributed with parameters $\sum_{i=1}^n \alpha_i$ and ρ . This is called the reproductive property of gamma distribution.
4. Find the 1, 2, 3 and 4 step transition probability matrix for the follow single step transition matrix. Does it exhibit any special characteristic?

$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

62 *System Reliability Modelling and Evaluation*

5. The state transition diagram of a continuous time Markov chain is given below. The states 1 and 2 are working states and state 3 is failed state. Calculate
- (a) The availability, i.e. the steady state probability of being in the working state
 - (b) MTTF
 - (c) Mean cycle time



References

1. U. Norayan Bhat, *Elements of Applied Stochastic Processes*, John Wiley (1972).
2. R. Billinton, *Power System Reliability Evaluation*, Gordon and Breach (1970).
3. J.A. Buzacott, Markov Approach to Finding Failure Times of Repairable Systems, *IEEE Transactions on Reliability*, **R-19**, 4, (1970).
4. A.B. Clarke and R.L. Disney, *Probability and Random Processes for Engineers and Scientists*, John Wiley (1970).
5. D.R. Cox and H.D. Miller, *The Theory of Stochastic Processes*, Methuen (1965).
6. W. Feller, *An Introduction to Probability Theory and Its Applications*, 1. (Third Edition) John Wiley (1958).
7. J.G. Kemeny and J.L. Snell, *Finite Markov Chains*, Van Nostrand (1960).
8. E. Parzen, *Stochastic Processes*, Holden-Day (1962).
9. G.H. Sandler, *System Reliability Engineering*, Prentice-Hall (1963).

CHAPTER 3

Frequency and Associated Concepts

Introduction

This chapter develops the frequency balancing approach to the stochastic process. This method has been used with considerable success in the field of power system reliability evaluation and is known as the frequency and duration approach. The concept of frequency is examined in detail and used to derive expressions for the mean cycle time and the mean duration of a state. In Chapter 5 the concept of frequency is used to derive conditions of mergeability which are very useful in dealing with large systems.

Interstate Transition Rate

Consider a large number of identical systems having the underlying stochastic process $Z(t)$. All the systems start in state i and further assume that in one step, state i can communicate with either state j or state k . If the histories of all the systems are now plotted they may appear as shown in Fig. 3.1. These observations are called independent realizations of the stochastic process $Z(t)$. Further, if N realizations in which the system transits to state j are separated out, the durations of the system in state i in this subset are called N independent realizations of the random variable X_{ij} , i.e. the duration of the state i under the condition that the system will transit to state j . In the stochastic process where all the states intercommunicate, these realizations could be obtained from a single long realization of the system by observing from the moment of each entry into state i until the termination of state i by transiting to state j .

Define:

$f_{ij}(x)$ = The probability density function of the random variable X_{ij}

$F_{ij}(x)$ = The distribution function of the random variable X_{ij}

$$= P(X_{ij} \leq x)$$

$$= \int_0^x f_{ij}(y) dy$$

and

$$\begin{aligned} S_{ij}(x) &= \text{The survivor function of the random variable } X_{ij} \\ &= 1 - F_{ij}(x) \\ &= \int_x^\infty f_{ij}(y) dy \end{aligned}$$

Now introduce the random variable

$$n(x) = \text{The number of realizations in which the system is in state } i \text{ at time } x.$$

Then

$$\begin{aligned} E\{n(x)\} &= \text{The expected value of } n(x) \\ &= N \int_x^\infty f_{ij}(y) dy \\ &= \eta(x) \end{aligned} \quad (3.1)$$

The expected number of transitions to state j in the interval $(x, x + \Delta x)$ is given by

$$\Delta\eta(x) = \eta(x) - \eta(x + \Delta x)$$

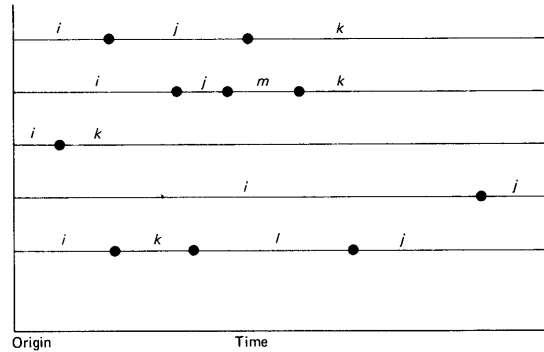


Fig. 3.1. Independent realizations of the stochastic process $Z(t)$.

and the rate of transition from state i to state j , per realization surviving up to x is

$$\begin{aligned} \lambda_{ij}(x) &= \frac{\Delta\eta(x)}{\eta(x)} \cdot \frac{1}{\Delta x} \\ &\quad \Delta x \rightarrow 0^+ \\ &= -\frac{1}{\eta(x)} \frac{d\eta(x)}{dx} \end{aligned} \quad (3.2)$$

Substituting the values of $\eta(x)$ and $\frac{d\eta(x)}{dx}$ from (3.1)

$$\lambda_{ij}(x) = \frac{f_{ij}(x)}{S_{ij}(x)} \quad (3.3)$$

The transition rate from state i to j is therefore the same as the hazard rate associated with the random variable X_{ij} . The above derivation is useful in getting an appreciation of the time specific transition rate as a relative expected rate. It is, however interesting to transform Equation (3.3) into another form.

$$\begin{aligned} \lambda_{ij}(x) &= \frac{f_{ij}(x)}{S_{ij}(x)} \\ &= \lim_{\Delta x \rightarrow 0^+} \frac{f_{ij}(x) \Delta x}{\Delta x} \frac{1}{S_{ij}(x)} \\ &= \lim_{\Delta x \rightarrow 0^+} \frac{P(x < X_{ij} \leq x + \Delta x)}{\Delta x} \frac{1}{P(x < X_{ij})} \\ &= \lim_{\Delta x \rightarrow 0^+} \frac{P(x < X_{ij} \leq x + \Delta x | x < X_{ij})}{\Delta x} \frac{P(x < X_{ij})}{P(x < X_{ij})} \\ &= \lim_{\Delta x \rightarrow 0^+} \frac{P(x < X_{ij} \leq x + \Delta x | x < X_{ij})}{\Delta x} \end{aligned} \quad (3.4)$$

That is, as $\Delta x \rightarrow 0^+$

$$\begin{aligned} \lambda_{ij}(x) \Delta x &= P(x < X_{ij} \leq x + \Delta x | x < X_{ij}) \\ &= \text{The probability of transiting from state } i \text{ to state } j \text{ in the} \\ &\quad \text{interval } (x, x + \Delta x) \text{ given that this transition has not taken} \\ &\quad \text{place up to time } x \\ &= \text{The probability of a single transition from state } i \text{ to state } j \\ &\quad \text{at the age } x \text{ of state } i \end{aligned}$$

Therefore, as $\Delta x \rightarrow 0^+$

The expected number of transitions from state i to state j in the interval $(x, x + \Delta x)$

$$\begin{aligned} &= 1 \cdot P(x < X_{ij} \leq x + \Delta x | x < X_{ij}) \\ &= \lambda_{ij}(x) \Delta x \end{aligned}$$

i.e. $\lambda_{ij}(x)$ = The expected or the mean transition rate from state i to state j at the age x of state i .

The quantity $\lambda_{ij}(x)$ is called the age specific transition rate and under particular conditions may be designated as the failure rate, hazard function, etc. The concept of the interstate transition rate will be further treated in the next chapter while dealing with the state transition diagram. When X_{ij} is exponentially distributed with probability density function $\rho \exp\{-\rho x\}$, the age specific transition rate, as shown in the last chapter, is constant, i.e.

$$\lambda_{ij}(x) = \lambda_{ij} = \rho = \frac{1}{\text{Mean value of } X_{ij}}$$

The transition rate is, therefore, constant, i.e. independent of the age of the state and is equal to the reciprocal of the mean of the random variable X_{ij} . This property is true only for the exponential distribution when all the random variables generating the stochastic process are exponentially distributed, the transition rates are constant and the process is Markovian. The case of constant transition rates will be treated in detail in this chapter. The discussion of non-Markovian processes is deferred to Chapter 6.

The Concept of Frequency

The state space X of the stochastic process $Z(t)$ is assumed to be partitioned into two disjoint subsets X^+ and X^- . If any state of the subset is entered, that subset is said to have been encountered.

Define:

$f_+(t)$ = The time specific frequency of encountering the subset X^+ . This is the expected rate at which X^+ is encountered at time t . For $\Delta t \rightarrow 0^+$, $f_+(t)\Delta t$ represents the expected number of times X^+ is encountered in the interval $(t, t + \Delta t)$.

$E_{ij}(t)$ = The time specific frequency of encountering state j from state i . This is the expected rate at which state j is encountered from state i or the mean rate at which the system transfers from state i to state j at time t .

$p_i(t)$ = The probability of being in state i at time t , for the given initial condition.

$p_+(t)$ = The probability of the system being in X^+ at time t for the given initial conditions.

$$= \sum_{i \in X^+} p_i(t)$$

λ_{ij} = The transition rate from state i to state j .

State j is encountered from state i at a rate λ_{ij} if the system is in state i , but if the system is not in state i then this rate is obviously zero. The transition rate or the encounter rate can therefore be represented by a discrete random variable $\beta_{ij}(t)$ such that

$$\begin{aligned} \beta_{ij}(t) &= \begin{cases} \lambda_{ij} & \text{if } Z(t) = i \\ 0 & \text{if } Z(t) \neq i \end{cases} \\ E_{ij}(t) &= \lambda_{ij}P\{Z(t) = i\} + 0 \cdot P\{Z(t) \neq i\} \\ &= \lambda_{ij}p_i(t) \end{aligned}$$

Since the states are all mutually exclusive

$$\begin{aligned} f_+(t) &= \sum_{i \in X^-} \sum_{j \in X^+} E_{ij}(t) \\ &= \sum_{i \in X^-} \sum_{j \in X^+} p_i(t) \lambda_{ij} \end{aligned} \quad (3.5)$$

If there is only one state j in X^+ , then the time specific frequency of encountering it is

$$f_j(t) = \sum_{i \neq j} p_i(t) \lambda_{ij} \quad (3.6)$$

Similarly, denoting the expected transition rate from j to $i \in X^-$ by $E_j(t)$

$$E_j(t) = p_j(t) \sum_{i \in X^-} \lambda_{ji}$$

As $\Delta t \rightarrow 0^+$, $f_j(t)\Delta t$ and $E_j(t)\Delta t$ represent respectively the expected number of transitions into and out of state j in the interval $(t, t + \Delta t)$. Since the

probability of more than one transition in $(t, t + \Delta t)$ can be reasonably assumed as $O(\Delta t)$, the quantities $f_j(t)\Delta t$ and $E_j(t)\Delta t$ can be interpreted as the probabilities of a single transition into and out of state j . The difference $[f_j(t)\Delta t - E_j(t)\Delta t]$ therefore represents an increase in the probability of being in state j , i.e.

$$\Delta p_j(t) = f_j(t)\Delta t - E_j(t)\Delta t$$

As $\Delta t \rightarrow 0^+$

$$\frac{dp_j(t)}{dt} = -p_j(t) \sum_{i \in X^-} \lambda_{ji} + \sum_{i \in X^+} p_i(t) \lambda_{ij} \quad (3.7)$$

This can be recognized as the forward differential equation of state j . In matrix form

$$Ap(t) = p'(t) \quad (3.8)$$

where

- $p(t)$ = The column matrix whose i th value $p_i(t)$ represents the probability of being in state i at time t for the given initial condition.
- $p'(t)$ = The differential of $p(t)$
- A = The transpose of the transition rate matrix used in the Markov approach.

The frequency balancing approach has been used to write the forward differential equations for the system and can be extended to derive expressions for the system reliability indices both in the time specific as well as the steady state domains.

Time Specific Domain

It is often necessary to examine the probable system performance over a finite interval of time. It is usual in the literature to define reliability indices in terms of system success or system failure. However, in many complex systems, there may be more than one degraded mode of failure. For example, a large chemical plant may not be just up or down but may have many possible capacity states. This is also true of transportation systems which may not be just available or unavailable and may have degrees of availability or unavailability. It is therefore appropriate to define reliability measures in terms of a subset X^+ where this subset may contain just one state or several states of the system. In particular applications the state will be referred to as success, failure, or some other appropriate name. In the transient domain the following indices are commonly used for repairable systems:

1 Time Specific Availability of Subset X^+

This is also designated in the literature as pointwise availability or instant availability and is the probability of the system being in any state contained in subset X^+ at a particular instant of time t . This will be denoted by $A_+(t)$ and since all states of the system are mutually exclusive, i.e. separated in time

$$A_+(t) = \sum_{i \in X^+} p_i(t) \quad (3.9)$$

The probability of being in state i at time t can be found from Equation (3.9) and has been discussed in Chapter 2. When X^+ is constituted by the system states which denote system failure, this can be called unavailability of the system at time t and designated by $U(t)$. This is probably the most widely used index in the transient domain.

2 Fractional Duration of Subset X^+

The fractional duration of subset X^+ in the interval (t_1, t_2) is also known as the interval availability or average time in X^+ and is defined as the expected proportion of the interval (t_1, t_2) spent in X^+ and denoted by $D_+(t_1, t_2)$. Since the states of the system are separated in time, the expected duration of X^+ is the sum of the expected durations in the states constituting X^+ , i.e.

$$D_+(t_1, t_2) = \frac{1}{t_2 - t_1} \sum_{i \in X^+} \int_{t_1}^{t_2} p_i(t) dt = \frac{\int_{t_1}^{t_2} A_+(t) dt}{t_2 - t_1} \quad (3.10)$$

Equation (3.10) can be understood by considering probability as a relative duration. The probability of being in X^+ at time t , i.e. $A_+(t)$ can be considered constant over the interval $(t, t + \Delta t)$, as $\Delta t \rightarrow 0^+$. Since $A_+(t)$ can be considered constant over the interval, in a very large number of realizations of the associated stochastic process, on the average $A_+(t)\Delta t$ time will be spent in X^+ during $(t, t + \Delta t)$. Considering (t_1, t_2) to be divided into m increments of Δt

$$\begin{aligned} D_+(t_1, t_2) &= \left[\sum_{m=1}^{\infty} A_+(t) \Delta t \right] / (t_2 - t_1) \\ &= \frac{\int_{t_1}^{t_2} A_+(t) dt}{t_2 - t_1} \end{aligned}$$

Another interpretation of Equation (3.10) can be provided by denoting the state occupied by t by $S_i(t)$ such that

$$\begin{aligned} S_i(t) &= 1 \text{ if the system is in state } i \\ &= 0 \text{ if otherwise.} \end{aligned}$$

The proportion of (t_1, t_2) spent in X^+ is, therefore

$$T_+(t_1, t_2) = \frac{1}{t_2 - t_1} \sum_{i \in X^+} \int_{t_1}^{t_2} S_i(t) dt$$

where the right hand side can be regarded as the limit of a time average taken at points $(0, \Delta t, 2\Delta t, 3\Delta t, \dots)$ as $\Delta t \rightarrow 0$. Therefore, the expected value $D_+(t_1, t_2) = E\{T_+(t_1, t_2)\}$

$$\begin{aligned} &= \frac{1}{t_2 - t_1} \sum_{i \in X^+} \int_{t_1}^{t_2} E\{S_i(t)\} dt = \frac{1}{t_2 - t_1} \sum_{i \in X^+} \int_{t_1}^{t_2} P\{S_i(t) = 1\} dt \\ &= \frac{1}{t_2 - t_1} \sum_{i \in X^+} \int_{t_1}^{t_2} p_i(t) dt = \frac{\int_{t_1}^{t_2} A_+(t) dt}{t_2 - t_1} \end{aligned}$$

3 Interval Frequency

The interval frequency $F_+(t_1, t_2)$ is defined as the expected number of times the subset X^+ is encountered in the interval (t_1, t_2) . Since the subset X^+ is said to have been encountered once, if the system transits from X^- to X^+ , $F_+(t_1, t_2)$ represents the expected number of transitions from X^- to X^+ and not from X^+ to X^- . In this treatment the state transition rates are assumed constant which puts the stochastic process in the Markovian class. Similar treatments can, however, also be made for non-Markovian processes.

The interval frequency can be obtained by integrating the Expression (3.5) for time specific frequency over the interval, i.e.

$$\begin{aligned} F_+(t_1, t_2) &= \int_{t_1}^{t_2} f_+(t) dt \\ &= \sum_{i \in X^-} \int_{t_1}^{t_2} p_i(t) \sum_{j \in X^+} \lambda_{ij} dt \end{aligned} \quad (3.11)$$

Sometimes the interval frequency may be divided by the interval length to obtain average interval frequency.

Methods of Calculation

When the state space of the system is small, it may be possible to find explicit expressions for the interval frequency and fractional duration. When the state space becomes relatively large, this approach is not feasible. Methods for obtaining time specific probabilities have already been described in Chapter 2. This section extends those methods to the calculation of interval frequency and fraction duration.

Method 1

The time specific state probabilities can be found by solving the following differential equation in matrix form

$$P'(t) = P(t)R$$

with the initial condition $P(0) = I$, i.e. the identity matrix.

Here, $P(t)$ = The matrix whose (ij) th term $p_{ij}(t)$ denotes the probability of being in state j given that the process was in state i at $t = 0$
 R = The transition rate matrix.

It was shown in the last chapter that if R has distinct eigenvalues then the probability matrix can be expressed as

$$P(t) = SD(t)S^{-1} \quad (3.12)$$

where

$D(t)$ = The diagonal matrix whose (ii) th element is $\exp(r_i t)$, r_i being the i th eigenvalue of R

S, S^{-1} = The matrices formed from the right and left eigenvectors of R

If the distribution at $t = 0$ is given by the row vector $p(0)$, the distribution of t is given by

$$p(t) = p(0)P(t)$$

The i th element of the row vector $p(t)$ is denoted by $p_i(t)$ and represents the probability of being in state i at time t for the given condition at $t = 0$. After finding $p_i(t)$, $A_+(t)$ can be calculated using Equation (3.9).

Fractional duration

Since only $D(t)$ on the right hand side of (3.12) is time dependent

$$\begin{aligned} \int_{t_1}^{t_2} p(t) dt &= p(0) S \int_{t_1}^{t_2} D(t) dt S^{-1} \\ &= p(0) S M(t_1, t_2) S^{-1} \end{aligned} \quad (3.13)$$

where

$$M(t_1, t_2) = \text{The diagonal matrix whose } (ii)^{\text{th}} \text{ term} = \frac{(e^{r_i t_2} - e^{r_i t_1})}{r_i}$$

Equation (3.13) thus yields $\int_{t_1}^{t_2} p_i(t) dt$ and by substituting these values in Equation (3.10), $D_+(t_1, t_2)$ can be calculated.

Interval Frequency

When the quantities $\int_{t_1}^{t_2} p_i(t) dt$ have been calculated by Equation (3.13), these can be substituted in (3.11) to determine $F_+(t_1, t_2)$

Method 2

If the initial state of the process is known, the differential equation becomes

$$p'(t) = p(t)R$$

The initial row vector $p(0)$ is such that its i th element $p_i(0)$ is equal to 1 if the process started in the i th state, otherwise it is 0. The process can be segmented into steps of a very small length Δt as shown in the last chapter and the probability vector at time $t = j\Delta t$ is

$$p(j\Delta t) = p((j-1)\Delta t) [I + R\Delta t] \quad (3.14)$$

The state probabilities can be calculated by repeated application of the above recursive relationship.

Fractional Duration

The fractional duration, using this technique, can be calculated using a discrete time equivalent of (3.10). Assuming $p_i(\Delta t)$ to be constant over the interval $(j\Delta t, (j+1)\Delta t)$

$$Dr(0, t) = \frac{1}{t} \sum_{j=0}^m p(j\Delta t) \Delta t \quad (3.15)$$

where

$Dr(0, t)$ = The row vector whose i^{th} entry, $D_i(0, t)$ represents the fractional duration of state i in the interval $(0, t)$.

Therefore

$$D_+(0, t) = \sum_{i \in X^+} D_i(0, t)$$

Interval Frequency

The interval frequency in the interval $(0, t)$ can be evaluated using a discrete time approximation of (3.11), i.e.

$$F_+(0, t) = t \sum_{i \in X^-} D_i(0, t) \sum_{j \in X^+} \lambda_{ij} \quad (3.16)$$

The above expression is readily understood by realizing that $t D_i(0, t)$ denotes the expected duration in state i , which when multiplied by the constant transition rate yields the expected number of transitions.

Steady State Domain

In many applications the time interval under consideration is very long, the stochastic process is remote from the time of origin and therefore the probability distribution has reached statistical equilibrium or is in a steady state condition. Under these conditions

$$\lim_{t \rightarrow \infty} p_i(t) = p_i$$

Therefore

$$\begin{aligned} E_{ij}(t) &= E_{ij} \\ &= \lambda_{ij} p_i \end{aligned}$$

Equation (3.5) in this limiting condition becomes

$$f_+ = \sum_{i \in X^-} \sum_{j \in X^+} p_i \lambda_{ij} \quad (3.17)$$

Now as $t \rightarrow \infty$ Equation (3.7) becomes

$$0 = -p_j \sum_{i \in X^-} \lambda_{ji} + \sum_{i \in X^-} p_i \lambda_{ij}$$

That is

$$\begin{aligned} f_j &= \sum_{i \in X^-} p_i \lambda_{ij} \\ &= p_j \sum_{i \in X^-} \lambda_{ji} \\ &= E_j \end{aligned} \quad (3.18)$$

Equation (3.18) describes the frequency balance of state j with the rest of the state space. This means that the frequency of encountering state j is equal to the frequency of encountering the rest of the state space from state j . The frequency of encountering a state may therefore be computed either by calculating the expected transition rate out of the state or into the state. Purely from a conceptual viewpoint, however, the frequency of encountering a state is the expected transition rate into the state. This definition holds in both the transient as well as the steady state domain. If the subset X^+ consists of more than one state, then by following the same reasoning as for (3.7)

$$\sum_{i \in X^+} \frac{dp_i(t)}{dt} = - \sum_{i \in X^+} p_i(t) \sum_{j \in X^-} \lambda_{ij} + \sum_{j \in X^-} p_j(t) \sum_{i \in X^+} \lambda_{ji}$$

As $t \rightarrow \infty$

$$\begin{aligned} f_+ &= \sum_{j \in X^-} p_j \sum_{i \in X^+} \lambda_{ji} \\ &= \sum_{i \in X^+} p_i \sum_{j \in X^-} \lambda_{ij} \\ &= f_- \end{aligned} \quad (3.19)$$

In further treatment, the steady state frequency will be simply denoted as frequency. Equations (3.17) and (3.19) are of fundamental importance in determining the frequency of cumulative or individual states.

The Frequency, the Cycle Time and the Mean Duration

The probability of being in X^+ at any time in the interval $(t, t + T)$ as $t \rightarrow \infty$ is $p_+ = \sum_{i \in X^+} p_i$ and the corresponding frequency is f_+ . Therefore, as $t \rightarrow \infty$, the interval frequency

$$\begin{aligned} F_+(t, t + T) &= \text{The expected number of times } X^+ \text{ is encountered in} \\ &\quad \text{the interval } (t, t + T). \\ &= Tf_+ \end{aligned}$$

Consequently

$$\begin{aligned} T^+ &= \text{The expected time between the two encounters of } X^+, \text{ i.e., the} \\ &\quad \text{mean cycle time} \\ &= T/F_+(t, t + T) \\ &= 1/f_+ \end{aligned} \quad (3.20)$$

Also

$$\begin{aligned} d_+ &= \text{The mean duration of } X^+, \text{ i.e., the expected time of stay in } X^+ \\ &\quad \text{in one cycle} \\ &= T^+ p_+ \end{aligned} \quad (3.21)$$

$$= p_+/f_+ \quad (3.22)$$

Equations (3.19), (3.20), and (3.22) are the backbone of the frequency and duration method of system reliability evaluation. The application of the concepts discussed is now illustrated with the help of the following example.

Example: The state space diagram of two independent and identical components is shown below in Fig. 3.2. The failure and repair of each component is denoted by λ and μ respectively. The state description is as follows:

- State 1 Both components up
- State 2 One component is up and the other is down
- State 3 Both components down

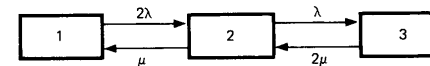


Fig. 3.2. State transition diagram of two independent identical components.

It is assumed that initially both the components are up. The calculation of various indices is now illustrated.

Time Specific Probabilities

The state differential equations can be written using the concept that the expected transition rate into the state minus the expected transition rate out of the state equals the rate of change of the state probability, i.e. using Equation (3.7)

$$\text{STATE 1: } p_1'(t) = -2\lambda p_1(t) + \mu p_2(t)$$

$$\text{STATE 2: } p_2'(t) = -(\mu + \lambda)p_2(t) + p_1(t)2\lambda + p_3(t)2\mu$$

$$\text{STATE 3: } p_3'(t) = -2\mu p_3(t) + p_2(t)\lambda$$

The Laplace transform of the above equations, using the initial condition

$$p_1(0) = 1, \quad p_2(0) = 0, \quad \text{and} \quad p_3(0) = 0$$

is as follows

$$s\bar{p}_1(s) = 1 - 2\lambda\bar{p}_1(s) + \mu\bar{p}_2(s)$$

$$s\bar{p}_2(s) = -(\mu + \lambda)\bar{p}_2(s) + \bar{p}_1(s)2\lambda + \bar{p}_3(s)2\mu$$

$$s\bar{p}_3(s) = -2\mu\bar{p}_3(s) + \bar{p}_2(s)\lambda$$

Solving the above equations

$$\bar{p}_1(s) = \frac{1}{s + 2\lambda} + \frac{2\lambda\mu(s + 2\mu)}{s(s + \lambda + \mu)(s + 2\lambda + 2\mu)(s + 2\lambda)}$$

$$\bar{p}_2(s) = \frac{2\lambda(s + 2\mu)}{s(s + \lambda + \mu)(s + 2\lambda + 2\mu)}$$

$$\bar{p}_3(s) = \frac{2\lambda^2}{s(s + \lambda + \mu)(s + 2\lambda + 2\mu)}$$

Expanding into partial fractions

$$\bar{p}_1(s) = \frac{1}{(\lambda + \mu)^2} \left[\frac{\mu^2}{s} + \frac{2\lambda\mu}{s + \lambda + \mu} + \frac{\lambda^2}{s + 2(\lambda + \mu)} \right]$$

$$\bar{p}_2(s) = \frac{2\lambda}{(\lambda + \mu)^2} \left[\frac{\mu}{s} + \frac{\lambda - \mu}{s + \lambda + \mu} - \frac{\lambda}{s + 2(\lambda + \mu)} \right]$$

$$\bar{p}_3(s) = \frac{\lambda^2}{(\lambda + \mu)^2} \left[\frac{1}{s} - \frac{2}{s + \lambda + \mu} + \frac{1}{s + 2(\lambda + \mu)} \right]$$

Converting the above

$$p_1(t) = \frac{1}{(\lambda + \mu)^2} [\mu^2 + 2\lambda\mu e^{-(\lambda + \mu)t} + \lambda^2 e^{-2(\lambda + \mu)t}]$$

$$p_2(t) = \frac{2\lambda}{(\lambda + \mu)^2} [\mu + (\lambda - \mu)e^{-(\lambda + \mu)t} - \lambda e^{-2(\lambda + \mu)t}]$$

$$p_3(t) = \frac{\lambda^2}{(\lambda + \mu)^2} [1 - 2e^{-(\lambda + \mu)t} + e^{-2(\lambda + \mu)t}]$$

It should be appreciated that since the two components are statistically independent, the above equations could have more easily been derived by finding the probabilities of each component and using the product rule of probabilities.

Steady State Probabilities

As $t \rightarrow \infty$, the exponential terms disappear. The steady state values are

$$p_1 = \frac{\mu^2}{(\lambda + \mu)^2}$$

$$p_2 = \frac{2\lambda\mu}{(\lambda + \mu)^2}$$

$$p_3 = \frac{\lambda^2}{(\lambda + \mu)^2}$$

The remaining indices are calculated for state 3. The calculation for the other states is left to the reader as an exercise.

Fractional Duration

$$\begin{aligned} D_3(0, T) &= \frac{1}{T} \int_0^T p_3(t) dt \\ &= \frac{\lambda^2}{T(\lambda + \mu)^2} \left[T - 2 \frac{1 - e^{-(\lambda + \mu)T}}{\lambda + \mu} + \frac{1 - e^{-2(\lambda + \mu)T}}{2(\lambda + \mu)} \right] \end{aligned}$$

As $T \rightarrow \infty$

$$\begin{aligned} D_3(0, T) &= \frac{\lambda^2}{(\lambda + \mu)^2} \\ &= p_3 \end{aligned}$$

Interval frequency

$$F_3(0, T) = \int_0^T p_2(t) \lambda dt$$

$$= \frac{2\lambda^2}{(\lambda + \mu)^2} \left[\mu T + \frac{\lambda - \mu}{\lambda + \mu} (1 - e^{-(\lambda + \mu)T}) - \frac{\lambda}{2(\lambda + \mu)} (1 - e^{-2(\lambda + \mu)T}) \right]$$

Steady State frequency

$$f_3 = p_3 2\mu$$

$$= \frac{2\lambda^2 \mu}{(\lambda + \mu)^2}$$

It can be seen that as $T \rightarrow \infty$, the average interval frequency approaches the steady state frequency.

$$\frac{F_3(0, t)}{T \rightarrow \infty} = \frac{2\lambda^2 \mu}{(\lambda + \mu)^2}$$

$$= f_3$$

Mean Cycle Time to the Encounter of State 3.

$$T_3 = \frac{1}{f_3}$$

$$= \frac{(\lambda + \mu)^2}{2\lambda^2 \mu}$$

Mean Duration of State 3.

$$d_3 = p_3 T_3$$

$$= \frac{1}{2\mu}$$

Frequency Equilibrium in a System of Independent Components

Equation (3.19) denotes the frequency equilibrium of the subset X^+ with the disjoint subset X^- . A case of special interest, from the point of view of the frequency equilibrium, is the system comprised of N independent binary components. The term 'binary components' is used here in the general sense of

a process which exists in either of the two states, up and down, the duration in each state being assumed exponentially distributed with the mean values of m and r . The reciprocal of m is λ which is the rate of transition from the up to the down state, and similarly μ , the reciprocal of r denotes the rate of transition from the down to the up state. In such a system it can be shown that in addition to frequency equilibrium of the state i with the disjoint subset containing the rest of the states, there is also a frequency equilibrium between any two individual states, i.e.

$$E_{ji} = \text{The frequency encounter of state } i \text{ from state } j, \text{ i.e. the expected transitions per unit time from state } j \text{ to state } i.$$

$$= E_{ij}$$

$$= \text{The frequency encounter of state } j \text{ from state } i.$$

This relationship is of considerable value when dealing with independent binary unit systems.

Proof: Let the state i comprise m components in the up state and k components in the down state. Whenever a component changes state, the system transits from state i to the state fitting the resulting description of component states. Let it be assumed that if the component '0' transits from the up to the down state, the system transits from state i to j . Then

$$E_{ij} = p_i \lambda_{ij}$$

$$= \left[\frac{\mu_0}{\mu_0 + \lambda_0} \prod_{p \in A} \frac{\mu_p}{\mu_p + \lambda_p} \prod_{q \in B} \frac{\lambda_q}{\mu_q + \lambda_q} \right] \lambda_0 \quad (3.23)$$

where A is a set containing all the components in the up state, except the component '0' which is in the up state but is not contained in the set A . Similarly B is a set containing all the components in the down state. Also

$$E_{ji} = p_j \lambda_{ji}$$

$$= \left[\frac{\lambda_0}{\mu_0 + \lambda_0} \prod_{p \in A} \frac{\mu_p}{\mu_p + \lambda_p} \prod_{q \in B} \frac{\lambda_q}{\mu_q + \lambda_q} \right] \mu_0 \quad (3.24)$$

From Equations (3.23) and (3.24)

$$E_{ij} = E_{ji}$$

Alternative Interpretation of Mean Cycle Time, Mean Duration and Mean Frequency

The emphasis in the preceding sections is on the frequency as an expectation of the state encounter rate. The mean cycle time is then obtained as the reciprocal of the mean frequency. This section describes an alternative approach by first deriving the expressions for mean times and then deducing the frequency relationships.

This treatment is intended for theoretically inclined readers and requires some acquaintance with renewal theory. This section may be omitted without any loss of continuity.

The entire state space is again assumed to be partitioned into two disjoint subsets X^+ and X^- .

Let

U = The random variable specifying the time of uninterpreted wandering of the system among the states $\{i \in X^+\}$, i.e. starting in $i \in X^+$, this is the time which the system spends in subset X^+ before once getting out.

D = The random variable defining the time of uninterrupted wandering of the system among the states $\{i \in X^-\}$.

The sequence of random variables U and D defines an alternating renewal process. This is shown in Fig. 3.3. The random variable $U + D$ specifies the cycle time, i.e. the time between two successive encounters of subset X^+ or X^- . The quantity of interest is the mean value T_m of the random variable $U + D$, i.e. $E(U+D) = E(U) + E(D)$ where

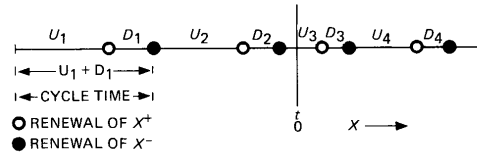


Fig. 3.3 The alternating renewal process (U, D) defined on the state space X .

$E(U)$ and $E(D)$ are respectively the mean values T_u, T_d of the random variables U and D . In long term or steady state analysis the interval of interest is $(t, t + T)$, $t \rightarrow \infty^+$. If the origin is, however, fixed at t , i.e. the time is represented by x such that $x = 0$, at t , then the origin of the process is at $t \rightarrow -\infty$, i.e. the process started remote from the time origin. Such a process is called an equilibrium alternating renewal process.

Let

$$\begin{aligned} S_U(u) &= \text{The survival function of } U \\ &= P\{U > u\} \end{aligned}$$

For the equilibrium renewal process, the residual life time of random variable U , i.e. the time y from a random instant $t \in U$ (i.e. under the condition that the system is in subset X^+ at the instant t) to the termination of U has the probability density function

$$f_{U_r}^e(x) = \frac{S_u(x)}{T_u}$$

The survival function $S_{U_r}(x)$ of the residual life time of U, U_r therefore is

$$S_{U_r}(x) = \frac{1}{T_u} \int_x^\infty S_U(u) du$$

Differentiating this expression

$$\frac{dS_{U_r}(x)}{dx} = -\frac{1}{T_u} S_U(x)$$

Since $S_U(0) = 1$

$$T_u = \left[-\frac{dS_{U_r}(x)}{dx} \Big|_{x=0} \right]^{-1} \quad (3.25)$$

Now

$$\begin{aligned} S_{U_r}(x) &= P\{U_r > x\} \\ &= \sum_{i \in X^+} P_i^* r_i(x) \end{aligned} \quad (3.26)$$

where

$P_i^* = P\{Z = i | i \in X^+\}$, i.e., the conditional steady state probability that the system is in the i th state given that it is in the subset X^+

$$Z = Z(t)$$

and

$r_i(x)$ = The probability that the system which begins to operate in state $i \in X^+$ at some random instant of time in the equilibrium process will not once get out of X^+ during x .

It is obvious that

$$P_i^* = p_i \left(\sum_{i \in X^+} p_i \right)^{-1} \quad (3.27)$$

Since the transition rates λ_{ik} are assumed constant, the transitions from state i to state k are governed by random variables which have exponential probability density functions. The probability density and the survival functions of the k th process will be

$$f_{ik}(x) = \lambda_{ik} e^{-\lambda_{ik}x}$$

and

$$S_{ik}(x) = e^{-\lambda_{ik}x}$$

When the probability density function is exponential, the instant of entry does not affect the probability density function, i.e. the probability density function of the residual life time of the random variable is the same as that of the random variable. Therefore, the probability that the system which begins to operate in the state $i \in X^+$ at some random instant of time, will not once get out of state i to state j

$$\begin{aligned} \eta_{ij}(x) &= \int_x^\infty f_{ij}(u) \prod_{\substack{k=1 \\ k \neq i, j}}^n S_{ik}(u) du \\ &= \int_x^\infty \lambda_{ij} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} u \right\} du \\ &= \lambda_{ij} \left(\sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} \right)^{-1} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \end{aligned}$$

Therefore

$$\begin{aligned} r_i(x) &= \sum_{j \in X^-} \eta_{ij}(x) + \sum_{j \in X^+} \int_0^x (1 - Q_j(y)) d\eta_{ij}(x - y) \\ &= \sum_{j \in X^-} \lambda_{ij} \left(\sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} \right)^{-1} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \\ &\quad + \sum_{j \in X^+} \lambda_{ij} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \int_0^x (1 - Q_j(y)) \exp \left\{ \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} y \right\} dy \end{aligned}$$

where $Q_j(y)$ is the probability distribution of the time at which the system gets out of X^+ given that the system entered state $j \in X^+$ at $y = 0$.

Substituting into (3.26)

$$\begin{aligned} S_{U_r}(x) &= \sum_{i \in X^+} P_i^* \left[\sum_{j \in X^-} \lambda_{ij} \left(\sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} \right)^{-1} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \right. \\ &\quad \left. + \sum_{j \in X^+} \lambda_{ij} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \int_0^x (1 - Q_j(y)) \right. \\ &\quad \left. \times \exp \left\{ \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} y \right\} dy \right] \\ \frac{dS_{U_r}(x)}{dx} &= \sum_{i \in X^+} P_i^* \left[\left(- \sum_{j \in X^-} \lambda_{ij} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \right) \right. \\ &\quad \left. + \sum_{j \in X^+} \left(\lambda_{ij} \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} ((1 - Q_j(x)) \exp \left\{ \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \right) \right. \\ &\quad \left. - 1 - \lambda_{ij} \left(\sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right) \exp \left\{ - \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} x \right\} \right. \\ &\quad \left. \times \int_0^x (1 - Q_j(y)) \exp \left\{ \sum_{\substack{k=1 \\ k \neq i}}^n \lambda_{ik} y \right\} dy \right] \end{aligned}$$

and

$$\left. \frac{dS_{U_r}(x)}{dx} \right|_{x=0} = - \sum_{i \in X^+} P_i^* \sum_{j \in X^-} \lambda_{ij}$$

Substituting into (3.25) and substituting the value of P_i^* from (3.27)

$$T_u = \sum_{i \in X^+} p_i \left(\sum_{i \in X^+} p_i \sum_{j \in X^-} \lambda_{ij} \right)^{-1} \quad (3.28)$$

Similarly

$$T_d = \sum_{i \in X^-} p_i \left(\sum_{i \in X^-} p_i \sum_{j \in X^+} \lambda_{ij} \right)^{-1} \quad (3.29)$$

The expressions for T_u and T_d have been derived by examining the distribution of the states constituting X^+ and X^- under stationary conditions. The behaviour of the alternating renewal process $\{U, D\}$ is now examined in detail. Since $x = 0$ is at $t \rightarrow \infty^+$ (see Fig. 3.3)

$P_u(x=0)$ = The probability that the system is in X^+ at the observation origin, given the process started a long time ago

$$= P_U(t \rightarrow \infty^+) \\ = \sum_{i \in X^+} p_i \\ = \frac{T_u}{T_u + T_d} = P_U(x) \quad (3.30)$$

Similarly

$$P_D(x=0) = P_D(x) = \sum_{i \in X^-} p_i = \frac{T_d}{T_u + T_d} \quad (3.31)$$

Substituting (3.30) and (3.31) into (3.28) and (3.29) respectively

$$T_m = T_u + T_d = \left(\sum_{i \in X^+} P_i \sum_{j \in X^-} \lambda_{ij} \right)^{-1} \\ = \left(\sum_{i \in X^-} p_i \sum_{j \in X^+} \lambda_{ij} \right)^{-1} \quad (3.32)$$

This expression for the cycle time is the same as that derived using the expectation concept. The frequency of encountering X^+ at x is the same as the renewal density of the random variable $U + D$. The renewal density is defined as

$$h(x) = \frac{E(N_{x,x+\Delta x})}{\Delta x} \\ \Delta x \rightarrow 0^+$$

where

$N_{x,x+\Delta x}$ = The random variable representing the number of renewals in $(x, x + \Delta x)$

E Denotes the expectation

Now, let

$h_{ij}(x)$ = The renewal density of subset j given that the system is in i at $x = 0$

It can be shown that for a modified renewal process, i.e. a sequence of independent random variables in which all the random variables are identically distributed except the first one which has a different distribution, the expression for renewal density is

$$\bar{h}(s) = \frac{\bar{f}_1(s)}{1 - \bar{f}(s)}$$

where

$\bar{f}_1(s)$, $\bar{f}(s)$ and $\bar{h}(s)$ are the Laplace transforms of $f_1(x)$, $f(x)$ and $h(x)$ respectively.

The quantities $f_1(x)$ and $f(x)$ are the probability density functions of the first and the subsequent random variables respectively.

1. Determination of $\bar{h}_{X^+X^-}(s)$

If the system is assumed to be in subset X^+ at 0, the first renewal of X^- or the encounter of X^+ occurs at $U_1 + D_1$ where the subscript indicates the number determined from $x = 0$ and not $t = 0$ (see Fig. 3.3) and the prime indicates that the distribution of the first U is different from the subsequent ones. The second renewal of X^- is at $U_2 + D_2$ and so on.

Now

$$f_{U_1}(x) = \frac{S_U(x)}{T_u}$$

and

$$f_{D_1}(x) = f_D(x)$$

Therefore

$$\bar{f}_1(s) = \frac{(1 - \bar{f}_U(s))\bar{f}_D(s)}{T_u \cdot s}$$

and

$$\bar{f}(s) = \bar{f}_U(s)\bar{f}_D(s)$$

Therefore

$$\bar{h}_{X^+X^-}(s) = \frac{(1 - \bar{f}_U(s))\bar{f}_D(s)}{T_u s (1 - \bar{f}_U(s)\bar{f}_D(s))} \quad (3.33)$$

2. Determination of $\bar{h}_{X^-X^-}(s)$

$$\bar{f}_1(s) = \frac{1 - \bar{f}_D(s)}{T_d s}$$

and

$$\bar{f}(s) = \bar{f}_U(s) \bar{f}_D(s)$$

Therefore

$$\bar{h}_{X^-X^-}(s) = \frac{1 - \bar{f}_D(s)}{T_d s (1 - \bar{f}_U(s) \bar{f}_D(s))} \quad (3.34)$$

Now

$$f_+(x) = h_{X^-}(x) = \bar{h}_{X^+X^-}(x) \cdot P\{Z(0) = i \in X^+\} \\ + h_{X^-X^-}(x) \cdot P\{Z(0) = i \in X^-\}$$

i.e.

$$\bar{f}_+(s) = \bar{h}_{X^-}(s) = \bar{h}_{X^+X^-}(s) \cdot P^+ + \bar{h}_{X^-X^-}(s) P^- \\ = \frac{1}{s(T_u + T_d)}$$

Converting

$$f_+(x) = f_+ = \frac{1}{T_u + T_d} \\ = \frac{1}{T_m} \\ = \sum_{i \in X^-} p_i \sum_{j \in X^+} \lambda_{ij} \\ = \sum_{i \in X^+} p_i \sum_{j \in X^-} \lambda_{ij} \\ = f_-(x) = f_- \quad (3.35)$$

since it can be proved similarly that

$$f_-(x) = f_- = \frac{1}{T_m}$$

Also from (3.29)

$$T_u = \text{The mean duration in } X^+ \\ = T_m P_+ \\ = P_+ / f_+ \quad (3.36)$$

It can, therefore, be seen that the expressions for frequency, mean cycle time and mean duration are the same as those derived using the expectation concept.

The Relationship to Average Values

It is quite well known that the arithmetic mean of a variable tends to its expected value as the number of trials becomes large. This section examines its application to the cycle time, the mean duration and the frequency indices. It can be seen from Fig. 3.3 that the cycle time between two successive encounters of X^+ is characterized by the random variable $T = U + D$. Thus $T_1 = U_1 + D_1$ is the cycle time to the first encounter of X^+ and $T_i = U_i + D_i$ is the cycle time between the $(i-1)$ th and the i th encounter of X^+ . The random variables T_i are independently and identically distributed with mean T_m . Assume the random variable T_i to be observed n times and define the random variable

$$\bar{T} = \frac{T_1 + T_2 + \dots + T_n}{n}$$

Then for any constant $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P\{|\bar{T} - T_m| > \epsilon\} = 0$$

This can be interpreted that, as the number of encounters of X^+ increases, the average cycle time approaches the mean cycle time with a probability of one. The mean cycle time found by Equation (3.32) is therefore the long run average interval between the two successive encounters of X^+ . Similarly it can be seen that T_u and T_d are the long run average residence periods of the system in X^+ and X^- . Since the frequency is the reciprocal of the mean cycle time, it can be appreciated that it is also a long term average.

The Concept of Equivalent Transition Rate

The concept of equivalent transition rate plays an important role in system reliability evaluation. Equation (3.5) can be written as

$$f_+(t) = \sum_{i \in X^-} \sum_{j \in X^+} p_i(t) \lambda_{ij} \\ = P_- \lambda_{X^-X^+}^{(e)}(t)$$

where

$$P_- = \sum_{i \in X^-} p_i(t)$$

and

$$\lambda_{X^-X^+}^{(e)}(t) = \text{The equivalent transition rate from subset } X^- \text{ to } X^+.$$

Therefore

$$\lambda_{X^-X^+}^{(e)}(t) = \sum_{i \in X^-} \sum_{j \in X^+} \lambda_{ij} / \sum_{i \in X^-} p_i(t) \quad (3.37)$$

The most important application of this concept is in reducing the system state space. The states can be merged and the equivalent transition rate from the merged states found by the application of Equation (3.37). The full implications of the equivalent transfer rate and the limitations on its use are outlined in Chapter 5 while deriving conditions of mergeability.

References

1. D.R. Cox, *Renewal Theory*, Methuen's Monographs on Applied Probability and Statistics, Methuen, London (1962).
2. D.R. Cox and H.D. Miller, *The Theory of Stochastic Processes*, Methuen, London (1965).
3. R.A. Howard, *Dynamic Probabilistic Systems*, John Wiley (1971).
4. C. Singh and R. Billinton, A Frequency and Duration Approach to Short Term Reliability Evaluation, *IEEE Trans.*, PAS-92, No. 6 (Nov./Dec. 1973).
5. C. Singh, 'Reliability Modelling and Evaluation in Electric Power Systems', Ph.D. thesis, University of Saskatchewan, Canada (1972).
6. I.A. Ushakov, Mean Time for Operation for a Semi-Markov Process, *Engrg. Cybernetics*, 4 (1969).

CHAPTER 4

System Reliability

Introduction

A system is a set of components arranged to accomplish a purpose or purposes under a given set of conditions. This chapter is concerned with determining the reliability characteristics of the system from the statistical information available on the failure and repair cycles of the components. Non-maintained systems are treated extensively in the literature and therefore the attention in this book is directed towards maintainable systems. In these systems, restorative action is initiated immediately after the failure and reliability modelling and evaluation is generally more complicated than in non-maintained cases. The theory and procedure is, however, quite general and can be equally applied to non-maintainable systems.

Numerical values of reliability measures can be obtained either through simulation or by solving mathematical models. The simulation approach is discussed in Chapter 7. This chapter describes various mathematical approaches to system reliability evaluation. The essential theory has already been outlined in the previous two chapters. Different approaches to system representation and solution are described. The following preliminary analysis is basic to all the approaches.

Definition and Description of the System and its Requirements

The first step in developing a reliability model is to define and describe the system and its requirements. The system must be categorized into major subsystems and the function of each subsystem and the interface between them defined. It is helpful to prepare a functional diagram showing the interaction of the various subsystems. A statement of the function of each subsystem and component should also be prepared. This type of information is easy to obtain for existing systems or systems where the design is in the final stage. At the early stages of a project, a good description of the system may not exist and the reliability engineer may have to come up with a reasonable description of the system from the study reports, development plans and specifications. Discussions with other personnel engaged on the project are very helpful. The system can be usually broken down into convenient blocks and most often the systems are designed on that basis. In terms of developing a model for the reliability

calculations, it may sometimes be preferable to group the components from one natural subsystem into another to create independent subsystems as this facilitates the reliability calculations.

Failure Modes and Effects Analysis

The analysis has to start at the level at which the information is available or can be made available. The term component is used in a general sense as it may in itself be a subsystem. For example, in generating capacity reliability analysis, a generating unit is regarded as a component whereas it is in itself a complex of components. The failure modes of the components should be investigated. The different modes are recognized from the different effects which they may have on the system. Some components may have just one mode of failure whereas others may have several. A relay or a circuit breaker may, for example, fail in an open or closed position. These modes generally have different effects on the system. A diode may for example fail as open circuited or short circuited and these failures develop into different system contingencies. After the different modes have been identified, their effects on the system should be studied. To assure complete systematic coverage of the effects of various failure modes, it is useful to employ some type of recording forms. It is again stressed that the identification of the failure modes is based on the different effects which these modes may have on the system. After the first analysis has been performed it may turn out that some of the effects, different though they may be, can be regarded as identical from the point of view of reliability analysis. It may therefore be possible to reduce the modes of failure. It is advisable to keep the failure modes to a minimum both from the point of view of reducing the complexity of the problem as well as from the point of view of limited statistical data. It is also sometimes possible to regard the different failure modes as independent without introducing any significant error. In such a case the component can be represented by a number of independent binary components equal to the failure modes.

The detailed form of above qualitative analysis is commonly referred to as FMEA (Failure Mode Effects Analysis). In certain programs which involve the limited production of complex and costly new hardware, it may not be possible to go much beyond this point and certain conclusions regarding design weaknesses may be derived from this type of analysis. Criticality analysis may be further added by ranking the critical failure mode effects according to the probability of occurrence. This is then called FMECA (Failure Mode, Effects and Criticality Analysis). This has been used extensively by NASA and other defence industries and is now finding its way into the analysis of commercial systems.

The purpose of the above two steps is to deepen systematically the understanding of the system in terms of the various component failure

contingencies. In some cases the system may be simple or so well known that it may be possible to by-pass this analysis because it is intuitively known to the reliability engineer. Assuming that the above analysis has been performed or the system is intuitively well understood, the techniques for mathematically deriving the reliability measures can be broadly classified as:

- i. state space approach
- ii. decomposition using conditional probability approach
- iii. network method

These techniques are now described in detail.

State Space Approach

A component may assume various states depending upon its failure and restorative modes. The system state describes the states of the components and the environment in which the system is operating. The set of all the possible states of the system is called the state space or event space. If the environment can exist in m states and the n components of the system are independent in each environment state, then the state space consists of 2^{n+m} states. The number of states is, however, modified because of the dependency restrictions. The state space approach involves the following steps:

- i. Enumerate all possible system states.
- ii. Determine the interstate transition rates. If a diagram is drawn showing the various states and the interstate transition rates, it is called the state transition diagram.
- iii. If the components are independent, the system state probabilities may be found from the component state probabilities by the product rule. In case of dependent failures, the system equations are formed and solved for the state probabilities.
- iv. The states are then grouped into subsets depending upon the requirements of the analysis. In most cases, measures are required only for success or failure but in some cases the indices may be computed for a graded mode of operation. After the grouping has been done, the subset probabilities, frequencies, cycle time and mean duration can be computed by the application of the formulae given in Chapter 3.

This approach is conceptually general and flexible and makes it possible to take into account various dependent failures. In large systems, especially involving dependent failures, it may be difficult to apply this technique. The methods for overcoming these difficulties are discussed in the next chapter. This approach is illustrated with the help of the following simple example.

System Description

The system consists of line 1 and line 2 supplying electric power to bus *C* from buses *A* and *B*. The supply at points *A* and *B* is considered to be perfectly reliable. Line 1 supplies 75% of the total power and line 2 supplies the remaining 25%. The lines can exist either in the 'up' or the 'down' state and the repair facilities are sufficient to perform repair on both the lines at the same time. The system is considered failed when it is supplying less than 75% of the power.

State Transition Diagram

In evolving the state transition diagram it is usually convenient to start from the state in which all the components are in a working condition. This condition is represented by state 1 in Fig. 4.1. The states to which the system can transit from state 1 are determined by the component transition modes and either of the lines can fail giving state 2 or 3. If the up times of both the lines are exponentially distributed with mean up times T_1 and T_2 , the transition rates from state 1 to states 2 or 3 are constant and given by

$$\lambda_1 = \frac{1}{T_1}, \quad \lambda_2 = \frac{1}{T_2}$$

This is shown in step 1.

Consider states 2 and 3 where one component is failed and the other is working. The system can change state either because of the failure of the working component or because of the repair of the failed component. If component 1 were to fail in state 2, the resulting state description would be (1D, 2D) as shown by state 4 in step 2 in Fig. 4.1. Since the failure rate λ_1 is constant for component number 1, it is unaffected by the time of residence in state 1 or state 2. If, however, the distribution were non-exponential, the transition rate would be dependent on times of residence and such a simple representation is not possible. This will be treated in detail in Chapter 6. Now considering state 3, if component 2 fails the system will transit to state 4 already generated but if component 1 is repaired, the system goes back to state 1. This is shown in step 2.

In state 4, both the components are down. Since both the components can be repaired independently, the resulting diagram is shown in step 3. The state transition diagram for this system is simple and could have been drawn in a straight-forward manner without going through the above procedure. The purpose of the above discussion is, however, to illustrate the concept of evolving the state transition diagram by examining the possible transition modes of each component in a particular system state. It should be realized that

it is not possible to draw a state transition diagram for relatively large systems. The above procedure is, however, readily programmable on the computer. The states in the computer program are represented by binary numbers, usually 0 standing for working components and 1 for failed ones. Special codes are employed for representing other modes (i.e. stand-by). The computer program usually evolves the state transition matrix directly and further manipulations are done using this matrix.

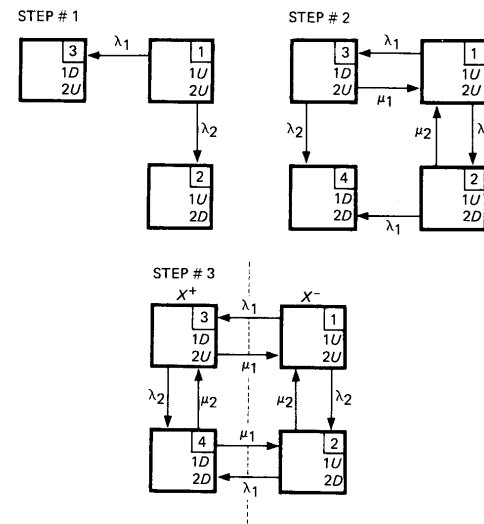


Fig. 4.1 Evolution of the state transition diagram.

State Probabilities

Both the lines are independent and the state probabilities can therefore be found by the simple product rule. It should be noted from Fig. 4.1 that when all the components are independent, each component can exert all its transition modes without any restriction from the other components or the environment. If in any system state, one or more of the transition modes of a component are suppressed, the components are no longer independent. Denoting the probabilities of the *i*th component being in the up or down state by $p_{iu}(t)$ and $p_{id}(t)$, these values for the initial condition $p_{iu}(0) = 1, p_{id}(0) = 0$ can be written from the results of the first chapter

$$p_{iu}(t) = \frac{\mu_i}{\lambda_i + \mu_i} + \frac{\lambda_i}{\lambda_i + \mu_i} e^{-(\lambda_i + \mu_i)t}$$

$$p_{id}(t) = \frac{\lambda_i}{\lambda_i + \mu_i} - \frac{\lambda_i}{\lambda_i + \mu_i} e^{-(\lambda_i + \mu_i)t}$$

The steady state values are

$$p_{iu} = \frac{\mu_i}{\lambda_i + \mu_i}$$

$$p_{id} = \frac{\lambda_i}{\lambda_i + \mu_i}$$

The state probabilities are

$$p_1(t) = p_{1u}(t)p_{2u}(t)$$

$$p_2(t) = p_{1u}(t)p_{2d}(t)$$

$$p_3(t) = p_{1d}(t)p_{2u}(t)$$

and

$$p_4(t) = p_{1d}(t)p_{2d}(t)$$

Reliability Measures

Since the system is considered failed if less than 75% of the total electric power is transferred, the subset X^+ of failed states is

$$X^+ = \{3, 4\}$$

and the disjoint subset X^- is

$$X^- = \{1, 2\}$$

The various measures can now be calculated

Time specific domain

The probability of the system being in the failed state at time t

$$= \text{The time specific availability of } X^+ = A_+(t)$$

$$= \sum_{i \in X^+} p_i(t)$$

$$\begin{aligned} &= p_3(t) + p_4(t) \\ &= p_{1d}(t)(p_{2u}(t) + p_{2d}(t)) \\ &= p_{1d}(t) \\ &= \frac{\lambda_1}{\lambda_1 + \mu_1} - \frac{\lambda_1}{\lambda_1 + \mu_1} e^{-(\lambda_1 + \mu_1)t} \end{aligned} \quad (4.1)$$

Fractional duration

$$\begin{aligned} &= D_+(0, T) \\ &= \frac{1}{T} \int_0^T A_+(t) dt \\ &= \frac{\lambda_1}{(\lambda_1 + \mu_1)T} \left[T + \frac{e^{-(\lambda_1 + \mu_1)T}}{\lambda_1 + \mu_1} - \frac{1}{\lambda_1 + \mu_1} \right] \\ &= \frac{\lambda_1}{\lambda_1 + \mu_1} \left[1 + \frac{1}{T(\lambda_1 + \mu_1)} (e^{-(\lambda_1 + \mu_1)T} - 1) \right] \end{aligned} \quad (4.2)$$

Interval frequency

$$\begin{aligned} &= F_+(0, T) \\ &= \int_0^T (p_1(t) + p_2(t)) \lambda_1 dt \\ &= \int_0^T p_{1u}(t) \lambda_1 dt \\ &= \frac{\lambda_1}{\lambda_1 + \mu_1} \int_0^T (\mu_1 + \lambda_1 e^{-(\lambda_1 + \mu_1)t}) dt \\ &= \frac{\lambda_1}{\lambda_1 + \mu_1} \left[\mu_1 T + \frac{\lambda_1}{\lambda_1 + \mu_1} (1 - e^{-(\lambda_1 + \mu_1)T}) \right] \end{aligned} \quad (4.3)$$

Steady State Analysis

The steady state measures can be obtained either by finding the limiting values of the time specific results or else by the direct application of the steady state formula. The latter approach is usually easier.

System unavailability

$$\begin{aligned}
 &= \text{The availability of } X^+ \\
 &= p_{1d} \\
 &= \frac{\lambda_1}{\lambda_1 + \mu_1}
 \end{aligned}$$

It should be noted that this result could also have been obtained either from Equation (4.1) by letting $t \rightarrow \infty$ or from Equation (4.2) as $T \rightarrow \infty$. The steady state unavailability can be interpreted, as mentioned before, either as the probability of being in the failed state at a point very remote from time of origin or as the limiting proportion of interval $(0, T)$ spent in the failed state, when T is very large. Frequency of encountering the failed state

$$\begin{aligned}
 &= (p_1 + p_2)\lambda_1 \\
 &= p_{1u}\lambda_1 \\
 &= \frac{\mu_1 \lambda_1}{\mu_1 + \lambda_1} \\
 &= p_{1d}\mu_1 \\
 &= \frac{\lambda_1 \mu_1}{\mu_1 + \lambda_1}
 \end{aligned}$$

This illustrates the frequency balance between the failed and working system state. The frequency could also be found from Equation (4.3)

by letting $T \rightarrow \infty$, in $F_+(0, T)/T$

Mean cycle Time

$$\begin{aligned}
 &= 1/\text{Frequency} \\
 &= \frac{\mu_1 + \lambda_1}{\mu_1 \lambda_1}
 \end{aligned}$$

Mean duration of the down state

$$\begin{aligned}
 &= \text{Mean cycle time} \times \text{unavailability} \\
 &= \frac{1}{\mu_1}
 \end{aligned}$$

The following section examines some basic configurations which occur quite commonly as systems or subsystems.

Series System

Components are considered to be in series when the failure of any one of the components causes system failure. This is shown in Fig. 4.2 where the failure of a component is equivalent to the removal of the corresponding block. It should be noted that the actual configuration of the components may or may not be in series, it is the effect of the failure of each component that is important. There are now two cases, firstly when all the components are independent and secondly



Fig. 4.2 Reliability diagram of n components in series

when after the system failure no further failure is possible so that there can be only one component in the failed state at any time. These two cases are considered separately.

Independent Case

Since the components are independent, the state probabilities can be derived from the component probabilities by the product rule. The state space will consist of $n+1$ subsets of states having $0, 1, \dots, n$ components in the failed state. The failure of even a single component causes system failure and therefore all the states except the one in which all the components are working represent system failure. It is therefore necessary to determine the measures for subset X^+ such that

$$\begin{aligned}
 X^- &= \{\text{state where all components are up}\} \\
 X^+ &= \{\text{all states except the one defined above}\} \\
 A_+(t) &= 1 - A_-(t) \\
 &= 1 - p_{1u}(t)p_{2u}(t) \dots p_{nu}(t)
 \end{aligned}$$

It can be seen that the above expression will involve exponential terms and the fractional duration and the interval frequency of the system failed state can be readily found by the application of appropriate formulae. In maintainable systems the steady state values are usually of primary importance and can be obtained from the following explicit formulae for this case.

System Unavailability

$$= 1 - p_{1u} p_{2u} \dots p_{nu}$$

$$= 1 - \frac{\mu_1 \mu_2 \dots \mu_n}{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2) \dots (\lambda_n + \mu_n)}$$

The frequency of encountering the down state

$$= A - \sum_{i=1}^n \lambda_i$$

$$= \frac{\mu_1 \mu_2 \dots \mu_n (\lambda_1 + \lambda_2 + \dots + \lambda_n)}{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2) \dots (\lambda_n + \mu_n)}$$

Mean cycle time to encounter the system down state

$$= \frac{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2) \dots (\lambda_n + \mu_n)}{\mu_1 \mu_2 \dots \mu_n (\lambda_1 + \lambda_2 + \dots + \lambda_n)}$$

Mean down time

$$= (\text{M.C.T.})(\text{Unavailability})$$

$$= \frac{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2) \dots (\lambda_n + \mu_n)}{\mu_1 \mu_2 \dots \mu_n (\lambda_1 + \lambda_2 + \dots + \lambda_n)} - \frac{1}{\lambda_1 + \lambda_2 + \dots + \lambda_n}$$

It should be noted that in the above expression the first term is the mean cycle time and the second term can be easily seen to be the mean up time. The mean down time is therefore represented as the difference of mean cycle time and the mean up time. For a two unit series system the expression for mean down time reduces to

$$\frac{1}{(\lambda_1 + \lambda_2)} \left[\frac{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2)}{\mu_1 \mu_2} - 1 \right]$$

$$= \frac{\lambda_1 \mu_2 + \lambda_2 \mu_1 + \lambda_1 \lambda_2}{(\lambda_1 + \lambda_2) \mu_1 \mu_2}$$

$$= \frac{\lambda_1 r_1 + \lambda_2 r_2 + \lambda_1 \lambda_2 r_1 r_2}{\lambda_1 + \lambda_2} \quad (4.4)$$

where r_1 and r_2 are the mean down times for components 1 and 2.

Dependent Failure

In many series systems it is safe to assume that once the system has failed, further failures will not occur. In such a case the state transition diagram is shown in Fig. 4.3. State 0 corresponds to the working state when all components are up. State $i, i=1, \dots, n$, corresponds to the i th component down. Once the system enters any of these states it can return only to the working state 0 because further failures are impossible. The transient solution can be obtained by solving the state differential equations. The steady state solution proceeds in the following manner. The frequency balance equations can be written as follows.

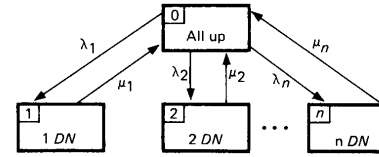


Fig. 4.3 A series system with dependent failure.

$$p_0 \sum_{j=1}^n \lambda_j = \sum_{i=1}^n p_i \mu_i \quad (4.5)$$

$$p_i \mu_i = p_0 \lambda_i \quad (4.6)$$

From Equation (4.6)

$$p_i = \frac{\lambda_i}{\mu_i} p_0$$

Substituting these values in equation

$$p_0 + \sum_{i=1}^n p_i = 1$$

$$p_0 = \frac{1}{Z}$$

where

$$Z = 1 + \sum_{i=1}^n \frac{\lambda_i}{\mu_i}$$

and

$$p_i = \frac{\lambda_i}{\mu_i Z}$$

System unavailability

$$= \sum_{i=1}^n p_i$$

$$= \frac{1}{Z} \sum_{i=1}^n \frac{\lambda_i}{\mu_i}$$

Frequency of encountering the down state

$$= p_0 \sum_{i=1}^n \lambda_i$$

$$= \frac{1}{Z} \sum_{i=1}^n \lambda_i$$

Mean cycle time

$$= Z / \sum_{i=1}^n \lambda_i$$

Mean Down Time

$$= (\text{M.C.T.})(\text{Unavailability})$$

$$= \sum_{i=1}^n \frac{\lambda_i}{\mu_i} / \sum_{i=1}^n \lambda_i$$

$$= \sum_{i=1}^n \lambda_i r_i / \sum_{i=1}^n \lambda_i \quad (4.7)$$

Referring to Equation (4.4), in high reliability component systems, the term $\lambda_1 \lambda_2 r_1 r_2$ can be neglected and therefore in such cases the mean down time is approximately the same in both the cases.

Series System with Spare

Consider the special case of a series system with a spare. The example chosen consists of a bank of three single phase transformers. As soon as a transformer fails, the entire bank is shut down until the failed unit has been replaced by the spare. The repair facilities are assumed to be unrestricted, i.e. as soon as the transformer fails, the repair is started irrespective of the fact that another unit may also be undergoing repair. The down time consists of two phases, the

repair time and the change out or the re-installation period. The following symbols are used:

λ = failure rate of a single phase transformer

μ = repair rate

γ = installation rate

The up time and the down times are assumed to be exponentially distributed. Reliability modelling of this system when the down times are not exponentially distributed is discussed in Chapter 6. The state transition diagram is shown in Fig. 4.4. The subsets X^+ , X^- representing the system failed and working states can be defined as follows:

$$X^+ = \{3, 4, 5\}$$

$$X^- = \{1, 2\}$$

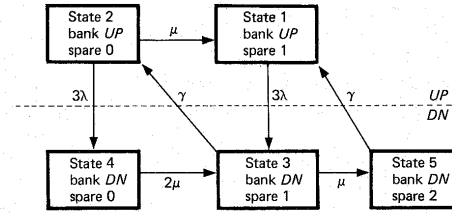


Fig. 4.4 The state transition diagram of a bank of three single phase transformers with one spare

The state frequency balance equations can be written as

$$3\lambda p_1 = \mu p_2 + \gamma p_5$$

$$(3\lambda + \mu)p_2 = \gamma p_3$$

$$(\gamma + \mu)p_3 = 3\lambda p_1 + 2\mu p_4$$

$$2\mu p_4 = 3\lambda p_2$$

$$\gamma p_5 = \mu p_3$$

Any four out of the above five equations together with the following equation can be used to obtain the steady state availabilities.

$$\sum_{i=1}^5 p_i = 1$$

The various probabilities obtained by solving these equations are:

$$p_1 = \frac{2\gamma^2\mu^2 + 2\gamma\mu^2(3\lambda + \mu)}{Z}$$

where

$$Z = 9\lambda^2\gamma^2 + 2\mu[(3\lambda + \mu)(3\lambda + \gamma)(\gamma + \mu)]$$

$$p_2 = \frac{6\lambda\gamma^2\mu}{Z} \quad p_3 = \frac{6\lambda\mu\gamma(3\lambda + \mu)}{Z}$$

$$p_4 = \frac{9\lambda^2\gamma^2}{Z} \quad p_5 = \frac{6\lambda\mu^2(3\lambda + \mu)}{Z}$$

$$p_{DN} = p_3 + p_4 + p_5 = \frac{9\lambda^2\gamma^2 + 6\lambda\mu\gamma(3\lambda + \mu) + 6\lambda\mu^2(3\lambda + \mu)}{Z}$$

and

$$p_{UP} = p_1 + p_2 = \frac{2\gamma^2\mu^2 + 2\gamma\mu^2(3\lambda + \mu) + 6\lambda\gamma^2\mu}{Z}$$

The frequency of encountering the subset X^+ of the failed states can be found as:

$$f_{X^+} = \sum_{\substack{m \in X^+ \\ k \in X^-}} p_m \lambda_{mk} = \sum_{\substack{k \in X^- \\ m \in X^+}} p_k \lambda_{km}$$

It is more convenient here to use the first relationship

$$\begin{aligned} f_{UP} = f_{DN} &= 3\lambda(p_1 + p_2) \\ &= 3\lambda p_{UP} \\ &= \frac{3\lambda(2\gamma^2\mu^2 + 2\gamma\mu^2(3\lambda + \mu) + 6\lambda\gamma^2\mu)}{Z} \end{aligned}$$

The mean cycle time

$$\begin{aligned} &= 1/f_{DN} \\ &= \frac{Z}{3\lambda(2\gamma^2\mu^2 + 2\gamma\mu^2(3\lambda + \mu) + 6\lambda\gamma^2\mu)} \end{aligned}$$

The mean down time

$$\begin{aligned} &= (\text{M.C.T.}) \times (\text{Unavailability}) \\ &= \frac{9\lambda^2\gamma^2 + 6\lambda\mu\gamma(3\lambda + \mu) + 6\lambda\mu^2(3\lambda + \gamma)}{3\lambda(2\gamma^2\mu^2 + 2\gamma\mu^2(3\lambda + \mu) + 6\lambda\gamma^2\mu)} \end{aligned}$$

Parallel Systems

When there are a number of components each one of them partially or completely serving the purpose, these components are said to be in parallel. When the components are simultaneously performing the same function so that the system will be fully available if at least one equipment is operating, it is said to be a parallel redundant configuration. In the example of two transmission lines of 25% and 75% of the total capacity requirements, this is a parallel but not a redundant arrangement. The following section deals with parallel redundant configurations. If the failure and repair rates of the equipment can be regarded as independent of the environment, the components can be usually regarded as independent. The probability of the system being in the down state at time t , i.e. the unavailability, can be found by the product rule

$$U(t) = p_{1d}(t)p_{2d}(t) \dots p_{nd}(t)$$

This product will be a function of exponential terms which are easily integrable. The fractional duration or the average down time in the interval $(0, t)$ can be therefore easily determined. For determining the interval frequency, the states in X^- from which the system can transit to X^+ in a single transition have to be identified. These states are called the boundary states and it is these states which contribute to interval frequency, the transitions from other states of X^- remaining within the subset. In the present example, these states are obviously those in which only one component is in the working state. The number of such states is obviously n . The time specific frequency of encountering the down state is

$$f_{DN}(t) = \sum_{i=1}^n p_{iu}(t) \lambda_i \prod_{\substack{j=1 \\ j \neq i}}^n p_{jd}(t)$$

and

$$f_{DN}(0, t) = \int_0^t f_{DN}(x) dx$$

The steady state results which are of main interest can be, however, more readily obtained in explicit form

$$U = \text{Unavailability} = \prod_{i=1}^n \frac{\lambda_i}{\lambda_i + \mu_i}$$

It has been explained previously that in the steady state condition, there is a frequency equilibrium between X^- and X^+ . The frequency can therefore be found for either of these two subsets. In practice, the one which is easiest to compute. In this case it is obvious that the easiest way is to find the frequency

of encountering the up state since there is only one state in the down state. Therefore

$$f_{DN} = f_{UP} = \left(\prod_{i=1}^n \frac{\lambda_i}{\lambda_i + \mu_i} \right) \sum_{i=1}^n \mu_i$$

The same result is derived for the sake of completeness as encounters from X^- to X^+ .

$$f_{DN} = \sum_{i=1}^n p_{iu} \lambda_i \prod_{\substack{j=1 \\ j \neq i}}^n p_{jd}$$

$$= \sum_{i=1}^n \mu_i \prod_{j=1}^n p_{jd}$$

$$= \left(\prod_{j=1}^n p_{jd} \right) \sum_{i=1}^n \mu_i$$

$$= \left(\prod_{i=1}^n \frac{\lambda_i}{\lambda_i + \mu_i} \right) \sum_{i=1}^n \mu_i$$

which is the same as derived above. The mean cycle time to encounter to the down state is the reciprocal of the frequency. The mean down time

$$= U/f_{DN}$$

$$= \frac{1}{\sum_{i=1}^n \mu_i}$$

The m/n Parallel System

In some systems of identical parallel components, m out of n components may be required for successful operation. These are termed m/n parallel systems and can be represented as shown in Fig. 4.5.

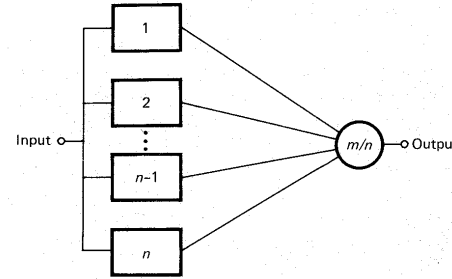


Fig. 4.5 The m/n Representation

If the components are independent, the reliability measures can be easily derived. Assuming identical failure and repair rates λ and μ

U = Unavailability

$$= \binom{n}{m-1} \left[\frac{\mu}{\lambda + \mu} \right]^{m-1} \left[\frac{\lambda}{\lambda + \mu} \right]^{n-m+1}$$

$$+ \binom{n}{m-2} \left[\frac{\mu}{\lambda + \mu} \right]^{m-2} \left[\frac{\lambda}{\lambda + \mu} \right]^{n-m+2} + \dots + \binom{n}{0} \left[\frac{\lambda}{\lambda + \mu} \right]^n$$

$$= \frac{1}{(\lambda + \mu)^n} \sum_{r=0}^{m-1} \binom{n}{r} \mu^r \lambda^{n-r}$$

For calculating the frequency of encountering the down state, the boundary states in X^- are those in which m units are working and $n-m$ are failed. Any further component failure results in system failure. The frequency is therefore

$$f_{DN} = \binom{n}{m} \frac{\mu^m \lambda^{n-m}}{(\lambda + \mu)^n} m \lambda$$

The mean cycle time can be found as a reciprocal of f_{DN} and the mean down time

$$= U/f_{DN}$$

Decomposition Using the Conditional Probability Approach

This method consists of decomposing a complex system into simple subsystems by the successive application of the conditional probability theorem. The idea is to calculate the reliability measures of the simpler subsystems and to combine these results to obtain the values for the system. The selection of the component or subsystem which is the key component or subsystem is therefore important. If this is not judiciously chosen, the final results will still be the same but the computation could be far more difficult. This method can be used to simplify both the state space as well as the network approach. Denoting the key component by X and representing the probability of its being up or down by $P(X)$ and $P(\bar{X})$ respectively

$$\begin{aligned} P_s &= \text{Probability of system success} \\ &= P(\text{System success} | X)P(X) + P(\text{System success} | \bar{X})P(\bar{X}) \quad (4.8) \end{aligned}$$

and

$$\begin{aligned} P_f &= \text{Probability of system failure} \\ &= P(\text{System failure} | X)P(X) + P(\text{System failure} | \bar{X})P(\bar{X}) \end{aligned}$$

The formula for calculating the frequency of encountering the down state (or any other state or subset of states) has been derived by the authors in the analysis of interconnected electric power systems and is explained below. The frequency of encountering the failed state of the system consists of the following components:

- i. the frequency of encountering the failed state, the component X being up, denoted by f_1
- ii. the frequency of encountering the failed state, the component X being down, denoted by f_2
- iii. the frequency of encountering the failed state as a result of the failure of component X , denoted by f_3

The frequency being the expectation of the state encounter rate

$$\begin{aligned} f_f &= \text{The frequency of encountering the system failure} \\ &= f_1 + f_2 + f_3 \\ &= f(\text{System failure} | X)P(X) + f(\text{System failure} | \bar{X})P(\bar{X}) + f_3 \end{aligned}$$

The expression for f_3 is now derived. Let the state space of the system be denoted by Y_0^- and Y_1^+ given that X is up or down respectively. Let the subsets of the failed and working states be denoted by superscripts + and - respectively. Y_0^- thus denotes the set of working states given X is up and Y_1^+ denotes the set of failed states given that X is down. States in Y_0^- and Y_1^+ having the same configuration of components, excluding X , are regarded as identical. It is assumed that the system cannot transit from the failed state to a working state by the failure of X . Therefore a state which is a member of Y_0^+ cannot be a member of

Y_1^- . The intersection of subsets Y_0^- and Y_1^+ represents the states in which the system is working if X is up but failed if X is down. Denoting these states by set S

$$f_3 = \sum_{i \in S} P(\text{State } i | X)P(X)\lambda_x$$

where

$$\lambda_x = \text{The failure rate of } X.$$

If the component X is independent of the rest of the system, then the sets Y_0 and Y_1 are equal and

$$P(\text{State } i | X) = P(\text{State } i | \bar{X})$$

Under this condition

$$\begin{aligned} \sum_{i \in S} P(\text{State } i | X) &= \sum_{i \in Y_1^+} P(\text{State } i | \bar{X}) - \sum_{i \in Y_1^+} P(\text{State } i | X) \\ &= P(\text{System failure} | \bar{X}) - P(\text{System failure} | X) \end{aligned}$$

The formula for system failure frequency, therefore, becomes

$$\begin{aligned} f_f &= f(\text{System failure} | X)P(X) + f(\text{System failure} | \bar{X})P(\bar{X}) \\ &\quad + (P(\text{System failure} | \bar{X}) - P(\text{System failure} | X))P(X)\lambda_x \quad (4.9) \end{aligned}$$

The application of this approach in simplifying reliability block diagrams is shown later in the chapter; the application in a state space approach is shown in the following example, which is rather an oversimplification of interconnected power systems.

Example 4.1: The electric power to an area is supplied from Station A where two identical generating units 1 and 2 are installed. A generating unit 3 is installed at a remote place and is supplying power to the same area through a transmission line. The failure and repair rates of the generating units are assumed to be λ and μ respectively.

The failure and repair rate of the transmission line are λ_x and μ_x . The system is classified as failed when no supply is available at all.

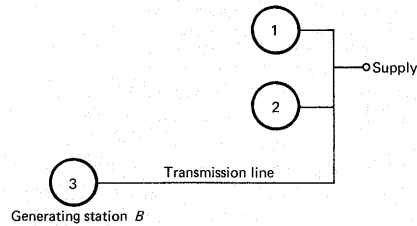


Fig. 4.6 Functional diagram of the generation scheme

Assuming the transmission line in the up state, the three units can be assumed as parallel, therefore

$$P(\text{System failure} | TL \text{ up}) = \left(\frac{\lambda}{\lambda + \mu} \right)^3$$

$$f(\text{System failure} | TL \text{ up}) = \left(\frac{\lambda}{\lambda + \mu} \right)^3 3\mu$$

Assuming the transmission line in the down state, the system is composed of only two parallel units, and therefore

$$P(\text{System failure} | TL \text{ dn}) = \left(\frac{\lambda}{\lambda + \mu} \right)^2$$

$$f(\text{System failure} | TL \text{ dn}) = \left(\frac{\lambda}{\lambda + \mu} \right)^2 2\mu$$

$$P_f = \left(\frac{\lambda}{\lambda + \mu} \right)^3 \frac{\mu_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu} \right)^2 \frac{\lambda_x}{\lambda_x + \mu_x}$$

$$= \frac{1}{\lambda_x + \mu_x} \left(\frac{\lambda}{\lambda + \mu} \right)^2 \left[\frac{\lambda \mu_x}{(\lambda + \mu)} + \lambda_x \right]$$

$$f_f = \left(\frac{\lambda}{\lambda + \mu} \right)^3 3\mu \frac{\mu_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu} \right)^2 2\mu \frac{\lambda_x}{\lambda_x + \mu_x}$$

$$+ \left[\left(\frac{\lambda}{\lambda + \mu} \right)^2 - \left(\frac{\lambda}{\lambda + \mu} \right)^3 \right] \frac{\mu_x \lambda_x}{\lambda_x + \mu_x}$$

$$= \left(\frac{\lambda}{\lambda + \mu} \right)^3 \frac{3\mu \mu_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu} \right)^2 \frac{2\mu \lambda_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu} \right)^2 \frac{\mu}{\lambda + \mu} \frac{\mu_x \lambda_x}{\lambda_x + \mu_x}$$

The same result can be obtained by drawing the state space diagram and making the calculations. This exercise is left to the reader for verification.

Network Approach

The state space approach is general and flexible but it becomes cumbersome when the number of states becomes large. The network approach when applicable usually provides a shorter route to solution. The network approach is usually not suitable when dependent failures or repairs are involved. Such failures occur in standby systems, interactive systems or systems whose failure and repair rates respond to a fluctuating common environment. It is not necessary to assume the event independence in this approach, but dependent events can greatly increase the algebra of computations and sometimes the solution may become impossible. At this point it is necessary to recognize the difference between two types of block diagrams.

Physical or Block Schematic Diagram

This diagram describes the actual connections between the components. Each block is a component and the diagram shows the manner in which they are actually connected.

Logic Diagram or Reliability Block Diagram

This diagram describes logical connections between components. Each block is a component which is removed when the component fails and replaced when it is repaired. The connections between the blocks describe the success or failure of the system as a function of the states of the component.

In the block schematic diagram, a component is not repeated because it represents the physical reality but in the reliability block diagram the block may be repeated. It is generally easy to construct the physical diagram as it follows from the physical layout. The reliability block diagram is usually difficult to prepare and in some cases a unique diagram may not exist. In the cases of information or power flow systems, it may be easy to construct a reliability block diagram. In many cases it may be difficult to do so. It requires a thorough understanding of the system and it is advisable to perform FMEA before constructing this block diagram. In the methods described below, the following assumptions are utilized.

1. the system is composed of independent components or subsystems. When the component reliabilities are high, this assumption enables approximate results to be obtained for non-independent units. Approximate formulae are also available for the reliability of a parallel network when the components respond to a two state fluctuating environment

2. each component or subsystem can be represented by a two state device, and the system success or failure can be expressed in terms of these two state devices
3. when all the components are working, the system is successful and when all the components are failed, the system is failed
4. when a group of components is working and the system is successful, the restoration of a failed component will not cause system failure
5. when a group of components is failed and the system is failed, the failure of any additional component will not restore the system to a successful state.

As the components are assumed to be independent it is not necessary to assume any particular distribution form for the up and down times in order to obtain steady state results. This will be clarified further in Chapter 6. It is only necessary to know the mean up and down times. The reciprocals of the mean up times and mean down times are failure and repair rates respectively. Assuming that the reliability block diagram exists, there are two main methods, the reliability diagram reduction method and a method based on manipulating the cut sets or tie sets. Both of these methods will now be described in detail.

Network Reduction Procedure

This method proceeds by the manipulation of the basic network structures:

- i. series structures
- ii. parallel structures
- iii. m/n structures when the n blocks originate from a common node

The method sequentially reduces the simple structures to equivalent units until the whole network reduces to a single unit. The necessity of assuming subsystem independence will be made clear in the next chapter while discussing the validity of equivalent transition rates in the context of subsystem reduction. Assuming that no subsystem is represented by more than one block, the procedure is as follows:

1. replace all series blocks by an equivalent block. In this case the following measures of the equivalent block can be easily obtained

$$\text{Availability} = \prod_{\forall i} p_{iu}$$

$$\text{Failure rate} = \sum_{\forall i} \lambda_i$$

2. in the resultant diagram replace the parallel and m/n substructures by equivalent blocks. In the parallel block diagrams it is convenient to determine the following measures of the equivalent block

$$\text{Unavailability} = \prod_{\forall i} p_{id}$$

$$\text{Repair rate} = \sum_{\forall i} \mu_i$$

The above steps are repeated until the whole network reduces to an equivalent block. If at any stage the network does not reduce any further, then the decomposition approach may be employed to generate simpler networks. In applying the decomposition approach to networks, the condition that the key component is good is equivalent to replacing this component by a short circuit and the condition that this component is down is equivalent to replacing it by an open circuit. This approach is illustrated by application to the system shown in Fig. 4.6. The reliability block diagram follows directly from the system topology and is shown in Fig. 4.7.

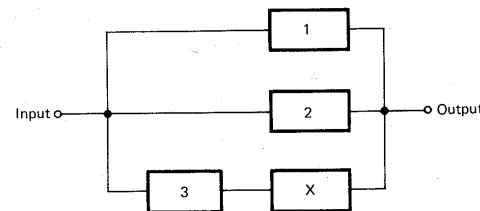


Fig. 4.7 Reliability block diagram of the system shown in Fig. 4.6

In this diagram, blocks 3 and X are in series. The equivalent block $3X$ has the following measures

$$p_{3xd} = 1 - p_{3u}p_{xu}$$

$$\lambda_{3x} = \lambda_3 + \lambda_x$$

Components 1, 2 and $3X$ are now in parallel, therefore

$$\text{Unavailability} = p_{1d}p_{2d}p_{3xd} = p_{1d}p_{2d} - p_{1d}p_{2d}p_{3u}p_{xu}$$

$$\begin{aligned}
&= p_{1d}p_{2d} - p_{1d}p_{2d}(1 - p_{3d})p_{xu} \\
&= p_{1d}p_{2d}(1 - p_{xu}) + p_{1d}p_{2d}p_{3d}p_{xu} \\
&= p_{1d}p_{2d}p_{xd} + p_{1d}p_{2d}p_{3d}p_{xu} \\
&= \left(\frac{\lambda}{\lambda + \mu}\right)^2 \frac{\lambda_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu}\right)^3 \frac{\mu_x}{\lambda_x + \mu_x}
\end{aligned}$$

This result is the same as obtained in Example 4.1. The repair rate of the block 3X can be derived from

$$\begin{aligned}
\frac{\lambda_{3x}}{\lambda_{3x} + \mu_{3x}} &= U_{3x} = 1 - p_{3u}p_{xu} \\
\text{i.e.} \quad \mu_{3x} &= \frac{\lambda_{3x}}{1 - p_{3u}p_{xu}} - \lambda_{3x} \\
&= \lambda_{3x}p_{3u}p_{xu}/(1 - p_{3u}p_{xu})
\end{aligned}$$

The frequency of system failure

$$\begin{aligned}
&= \text{System Unavailability} \sum_i \mu_i \\
&= p_{1d}p_{2d}(1 - p_{3u}p_{xu})(\mu_1 + \mu_2 + \mu_{3x}) \\
&= 2\mu \left[\left(\frac{\lambda}{\lambda + \mu}\right)^2 \frac{\lambda_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu}\right)^3 \frac{\mu_x}{\lambda_x + \mu_x} \right] \\
&\quad + p_{1d}p_{2d}(\lambda_3 + \lambda_x)p_{3u}p_{xu} \\
&= 2\mu \left[\left(\frac{\lambda}{\lambda + \mu}\right)^2 \frac{\lambda_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu}\right)^3 \frac{\mu_x}{\lambda_x + \mu_x} \right] \\
&\quad + \left(\frac{\lambda}{\lambda + \mu}\right)^2 (\lambda + \lambda_x) \frac{\mu}{\lambda + \mu} \frac{\mu_x}{\lambda_x + \mu_x} \\
&= \left(\frac{\lambda}{\lambda + \mu}\right)^3 \frac{3\mu\mu_x}{\lambda_x + \mu_x} + \left(\frac{\lambda}{\lambda + \mu}\right)^2 \frac{2\mu\lambda_x}{\lambda_x + \mu_x} \\
&\quad + \left(\frac{\lambda}{\lambda + \mu}\right)^2 \frac{\mu}{\lambda + \mu} \frac{\mu_x}{\lambda_x + \mu_x}
\end{aligned}$$

which is the same as derived in Example 4.1. It should be noted that though the deriving of explicit expressions could be a formidable task, the numerical results even for large systems can be obtained easily using a calculator or a computer.

Example 4.2: The reliability block diagram of a system is shown in Fig. 4.8. The mean up times and down times of the various blocks are:

$$\begin{aligned}
MUT_1 &= MUT_2 = MUT_3 = MUT_4 = MUT_5 = \frac{1}{2} \text{ Yr} \\
MDT_1 &= MDT_2 = MDT_3 = MDT_4 = MDT_5 = 20 \text{ Hr}
\end{aligned}$$

Assuming that all components are statistically independent, calculate the reliability measures for continuity between s and l .

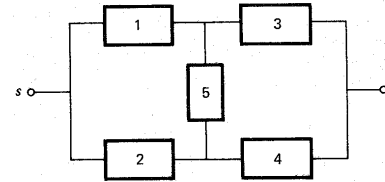


Fig. 4.8 Reliability block diagram for Example 4.2

It is obvious that no simple series or parallel paths exist in this reliability structure. Applying the principle of decomposition, the block diagram can be split into simple series and parallel configurations as shown in Fig. 4.9 in the form of a tree. Referring to Formulae (4.8) and (4.9), the values required are the probability and frequency of system failure given 5 is good and then given 5 is bad. The system measures can then be calculated in terms of these values.

The failure and repair rates of the components are:

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda$$

and

$$\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu$$

The probability of a component being down

$$p_d = \frac{\lambda}{\lambda + \mu}$$

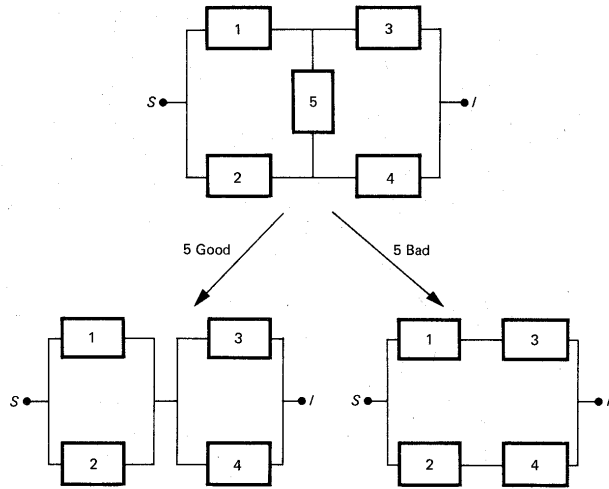


Fig. 4.9 Decomposition of a complex reliability block diagram into simple structures

Given 5 Good

$$p_{12d} = p_d^2$$

$$p_{12u} = 1 - p_{12d} = 1 - p_d^2$$

$$\begin{aligned} f_{12} &= p_{12d}(\mu_1 + \mu_2) \\ &= 2\mu p_d^2 \end{aligned}$$

The equivalent failure rate of component 12 is given by (3.37)

$$\begin{aligned} \lambda_{12} &= f_{12}/p_{12u} \\ &= 2\mu p_d^2/(1 - p_d^2) \end{aligned}$$

and the equivalent repair rate

$$\mu_{12} = f_{12}/p_{12d} = 2\mu$$

The values for the equivalent block 34 are the same as for 12 because they consist of identical components. The equivalent blocks 12 and 34 are in series and therefore

$$\begin{aligned} P(\text{System failure} | 5 \text{ Good}) &= p_{1234d} \\ &= 1 - p_{12u}p_{34u} \\ &= 1 - (1 - p_d^2)^2 = 2p_d^2 - p_d^4 \\ f(\text{System failure} | 5 \text{ Good}) &= p_{1234u}(\lambda_{12} + \lambda_{34}) \\ &= 4\mu p_d^2(1 - p_d^2) \end{aligned}$$

Given 5 Bad

$$p_{13u} = p_{1u}p_{3u} = (1 - p_d)^2$$

$$\begin{aligned} p_{13d} &= 1 - p_{13u} \\ &= 2p_d - p_d^2 \end{aligned}$$

$$\begin{aligned} f_{13} &= p_{13u}(\lambda_1 + \lambda_3) \\ &= (1 - p_d)^2 2\lambda \end{aligned}$$

$$\begin{aligned} \mu_{13} &= f_{13}/p_{13d} \\ &= \frac{2\lambda(1 - p_d)^2}{2p_d - p_d^2} \end{aligned}$$

$P(\text{System failure} | 5 \text{ Bad})$

$$\begin{aligned} &= p_{1324d} = p_{13d}p_{24d} \\ &= (2p_d - p_d^2)^2 \end{aligned}$$

$f(\text{System failure} | 5 \text{ Bad})$

$$\begin{aligned} &= p_{1324d}(\mu_{13} + \mu_{24}) \\ &= (2p_d - p_d^2)^2 \frac{4\lambda(1 - p_d)^2}{2p_d - p_d^2} \\ &= 4\mu p_d(1 - p_d)(2p_d - p_d^2) \end{aligned}$$

These results can now be combined using formulae (4.8) and (4.9).

$$\begin{aligned} P_f &= P(\text{System failure} | 5 \text{ Good})P(5 \text{ good}) \\ &\quad + P(\text{System failure} | 5 \text{ Bad})P(5 \text{ Bad}) \\ &= (2p_d^2 - p_d^4)(1 - p_d) + (2p_d - p_d^2)^2 p_d \end{aligned}$$

$$= 2p_d^2 + 2p_d^3 - 5p_d^4 + 2p_d^5$$

and

$$\begin{aligned} f_f &= f(\text{System failure} | 5 \text{ Good})P(5 \text{ Good}) \\ &\quad + f(\text{System failure} | 5 \text{ Bad})P(5 \text{ Bad}) \\ &\quad + (P(\text{System failure} | 5 \text{ Bad}) \\ &\quad - P(\text{System failure} | 5 \text{ Good}))P(5 \text{ Good}) \cdot \lambda_s \\ &= 4\mu p_d^2(1 - p_d^2)(1 - p_d) + 4\mu p_d^2(2p_d - p_d^2)(1 - p_d) \\ &\quad + ((2p_d - p_d^2)^2 - 2p_d^2 + p_d^4)\mu p_d \\ &= \mu(4p_d^2 + 6p_d^3 - 20p_d^4 + 10p_d^5) \end{aligned}$$

$$\text{Mean cycle time} = 1/f_f$$

and

$$\text{Mean Down Time} = p_f/f_f$$

Cut Set or Tie Set Methods

The network reduction approach is quite useful when the block diagram consists essentially of series and parallel structures. When the reliability block diagram is complex, decomposition into simple series and parallel paths may not be easy. The process could be quite difficult to program because it would require a lot of scanning. The approach using cut sets or tie sets is especially useful for computer applications. The following definitions are useful in appreciating this approach.

Simple Path

If in going from node x to y no node is traversed more than once, the path is simple. For example, in Fig. 4.8, 1–5–4 is a simple path.

Connected Subnetwork

A subnetwork is said to be connected if there exists a simple path between all pair of nodes.

Cut Set

This is a set of components whose failure alone will cause system failure. A minimal cut has no proper subset of components whose failure alone will cause system failure.

The minimal cuts for the reliability block diagram shown in Fig. 4.8 are listed in Table 4.1.

Table 4.1 Minimal cuts between input and output nodes in Fig. 4.8

Minimal cut set	Components in the set
C_1	1, 2
C_2	3, 4
C_3	1, 4, 5
C_4	2, 3, 6

Path or Tie Set

This is a set of components whose functioning alone will guarantee system success. A minimal path or tie set has no proper subset of components whose functioning alone would ensure system success.

In this approach it is possible to proceed either through cut set manipulation or tie set manipulation. The choice is usually dictated by their relative numbers. Both methods are explained.

Tie Set Manipulation

In the minimal path all the blocks constituting it are in series. The failure of any one of these blocks would render that path ineffective. The minimal paths themselves are, however, in parallel as the system will be successful so long as there is even one path available between the input and output of the reliability block diagram. Denoting the path available and unavailable by T and $\bar{T}$ respectively

$$P_s = P(T_1 \cup T_2 \cup T_3 \cup \dots \cup T_m)$$

$$= [P(T_1) + P(T_2) + P(T_3) + \dots + P(T_m)] \leftarrow \binom{n}{1} \text{ terms}$$

$$- [P(T_1 \cap T_2) + P(T_1 \cap T_3) + \dots + P(T_i \cap T_j)] \leftarrow \binom{n}{2} \text{ terms}$$

$$+ [P(T_1 \cap T_2 \cap T_3) + P(T_1 \cap T_2 \cap T_4) + \dots$$

$$+ P(T_i \cap T_j \cap T_k)] \leftarrow \binom{n}{3} \text{ terms}$$

...

$$+ (-1)^{n-1} [P(T_1 \cap T_2 \cap \dots \cap T_n)] \leftarrow \binom{n}{n} \text{ terms} \quad (4.10)$$

The total number of terms in this expression is $2^n - 1$ where n is the number of tie sets. It can be seen from Equation (4.10) that the independence of components need not be postulated. All that is needed is the evaluation of all the terms of the expansion (4.10). In the case of dependent failures, however, the entire state transition diagram may have to be drawn to evaluate the above terms. This method is therefore of importance for independent components. As the number of tie sets, however, increases, the expansion of all the terms becomes a formidable task. In such cases useful approximate formulae can be obtained by Boole's inequality.

$$P(T_1 \cup T_2 \cup \dots \cup T_n) \leq P(T_1) + P(T_2) + P(T_3) + \dots + P(T_n)$$

Therefore if only the first row in the expansion (4.10) is calculated, the result will be an upper bound approximation. This upper bound becomes a good approximation when the component reliabilities are low. As the components are assumed independent

$$P(T_i) = \prod_{j \in T_i} p_{ju}$$

When the number of tie sets is relatively small so that explicit relationships can be derived, the following procedure can be used to obtain exact values:

1. Calculate the tie set availabilities

$$P(T_i) = \prod_{j \in T_i} p_{ju}$$

2. The formula for system availability assuming the paths are independent is

$$P_s = 1 - \prod_{i=1}^n (1 - p(T_i))$$

3. Independence of paths has been assumed in the above expression. It will contain some terms containing p_{ju}^m . These terms are introduced due to the path independence assumption in calculating the 2nd to n th row in expansion (4.10). Therefore, the exact result for P_s may be obtained by replacing p_{ju}^m by p_{ju} . This procedure is, however, possible only when the number of tie sets is small and an explicit expression can be easily derived. When only the numerical results are to be obtained, either the expansion (4.10) may be evaluated or the upper bound approximation may be calculated.

Cut Set Manipulation

In the minimal cut the blocks are in parallel as all of them must fail, to produce a cut. The minimal cuts themselves are, however, in series as even a single cut ensures failure. Denoting the failure of the i th cut set by $\bar{C}_i$, the probability of system failure is

$$P_f = P(\bar{C}_1 \cup \bar{C}_2 \cup \bar{C}_3 \cup \dots \cup \bar{C}_m) \quad (4.11)$$

and

$$P_s = 1 - P_f$$

The Expression (4.11) can be expanded in the same manner as (4.10) and all the comments on (4.10) apply equally well to this expansion. By Boole's inequality

$$P_f \leq P(\bar{C}_1) + P(\bar{C}_2) + \dots + P(\bar{C}_m)$$

This upper bound of the probability of failure is useful in the high reliability region. The above inequality can be manipulated into the following form

$$P_s \geq 1 - [P(\bar{C}_1) + P(\bar{C}_2) + \dots + P(\bar{C}_m)]$$

This lower bound is good when the component reliabilities are high. The components being assumed independent

$$P(C_i) = 1 - \prod_{j \in C_i} p_{jd}$$

When the explicit formulae can be derived, the following procedure can be used to obtain the exact values.

1. Calculate the cut set availabilities

$$P(C_i) = 1 - \prod_{j \in C_i} p_{jd}$$

2. The formula for system success assuming the cut set independence is

$$P_s = \prod_{i=1}^m P(C_i)$$

3. Replace p_{ju}^m by p_{ju} , the resulting formula will be the exact expression for system reliability.

Frequency Calculation Using The Cut Set Approach

In order to understand the derivation of the failure frequency formula, the relationship between the minimal cut set and the system state-space should be understood. Consider a minimal cut set C_i which has components l and m as its members. This means that if the components l and m fail, the system will be failed irrespective of the states of the other components of the system. The failure of the members of C_i is equivalent to the system being in subset S_i of the state space S , where

$$S_i = \{s_j: \text{in the state } s_j, \text{ the components } l \text{ and } m \text{ are failed and the other components exist in a particular state}\}$$

The state s_j^v in which l and m are failed and all of the other components are functioning is called the vertex state of the subset S_i . The system can transit from the vertex state either upwards (in the sense of less components in the failed state) by the repair of the failed components l and m or it can transit downwards (in the sense of more components in the failed state) by the successive failures of more components. The subset S_i is constituted by the states generated by the downward transitions from s_j^v . The system could transit out of S_i by the repair of l or m and therefore the frequency of encountering subset S_i is

$$\begin{aligned} f_i &= \sum_{s_j \in S_i} P(s_j) \sum_{k \in C_i} \mu_k \\ &= \left(\sum_{k \in C_i} \mu_k \right) \left(\sum_{s_j \in S_i} P(s_j) \right) \\ &= \left(\sum_{k \in C_i} \mu_k \right) (P(S_i)) \\ &= \left(\sum_{k \in C_i} \mu_k \right) P(\bar{C}_i) \\ &= \bar{\mu}_i P(\bar{C}_i) \end{aligned}$$

where

$$\bar{\mu}_i = \sum_{k \in C_i} \mu_k$$

The relationship between the cut set and its equivalent state space subset can be more clearly understood by reference to Fig. 4.10 where the cut set C_1 of Fig. 4.8 and the equivalent subset S_1 are shown. $S_1 = \{s_1, s_2, \dots, s_8\}$. The states which are members of S_1 are generated by successive failures of components 3, 4 and 5 from the vertex state s_1 . From any state eS_1 , the system

could transit out of S_1 by the repair of component 1 or 2 which are the members of C_1 . The frequency of encountering S_1 is

$$\begin{aligned} f_1 &= (\mu_1 + \mu_2) \left[\sum_{i=1}^8 P(s_i) \right] \\ &= (\mu_1 + \mu_2) P(\bar{C}_1) \\ &= (\mu_1 + \mu_2)(p_{1d}p_{2d}) \end{aligned}$$

Now consider another minimal cut set C_k and its equivalent subset S_k of the state space S . The reader can appreciate the arguments by taking $C_k \equiv C_3$ in Fig. 4.10 where C_3 is the minimal cut set of Fig. 4.8. If S_i and S_k were mutually exclusive, there could not be any transition between S_i and S_k . This can be seen as follows. Suppose that from state s_i^v of subset S_i , a transition is possible to state s_u^k of subset S_k by the failure of a component, then the state s_u^k and all the states generated from s_u^k by the downward transitions will be common to both S_i and S_k . The two subsets will not therefore remain mutually exclusive. Reasoning backwards, there cannot be transitions between two mutually exclusive subsets equivalent to two minimal cut sets. In such a case the frequency contribution due to S_i and S_k is $f_i + f_k$. In practice, however, the state space subsets representing minimal cutsets overlap and the frequency formula for this case can be derived by referring to the Venn diagram in Fig. 4.11.

Let

$$S_i = A_1 + A_2$$

and

$$S_k = A_3 + A_2$$

so that

$$S_i \cap S_k = A_2$$

The frequency of encountering the subset $S_i \cup S_k$ is given by

$$f(S_i \cup S_k) = P(A_1)\bar{\mu}_i + P(A_3)\bar{\mu}_k + P(A_2)(\bar{\mu}_i \cap \bar{\mu}_k) \quad (4.12)$$

where

$$\bar{\mu}_i = \sum_{j \in C_i} \mu_j$$

$$\bar{\mu}_k = \sum_{j \in C_k} \mu_j$$

and

$$\bar{\mu}_i \cap \bar{\mu}_k = \sum_{j \in C_i \cap C_k} \mu_j$$

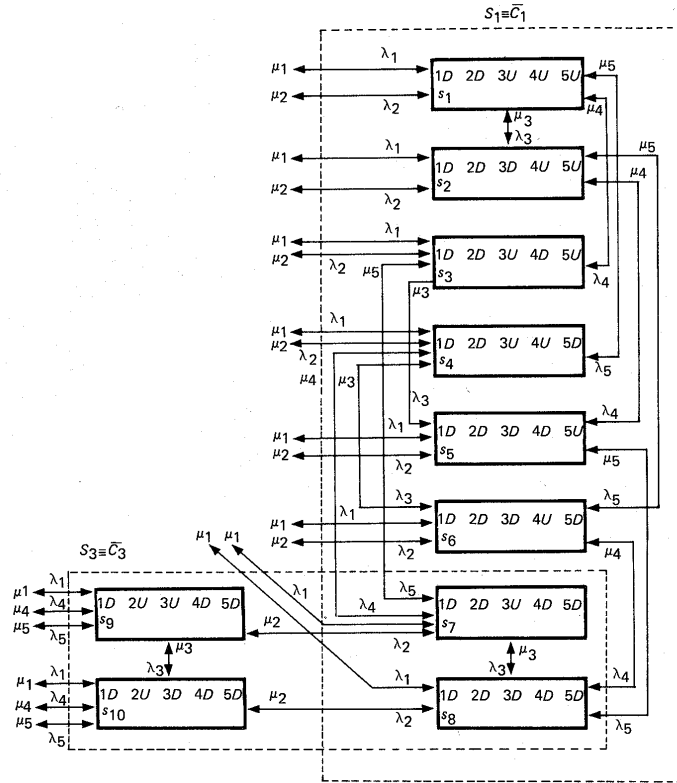


Fig. 4.10 The equivalence between minimal cut sets and state space representations

That is $\bar{\mu}_i \cap \bar{\mu}_k$ represents the summation of the repair rates of the components common to both minimal cut sets. For example for C_1 and C_3 , $\bar{\mu}_1 \cap \bar{\mu}_3 = \mu_1$ since only component 1 is common to both cut sets.

Equation (4.12) can be written as

$$\begin{aligned} f(S_i \cup S_k) &= P(A_1 + A_2)\bar{\mu}_i + P(A_3 + A_2)\bar{\mu}_k + P(A_2)(\bar{\mu}_i \cap \bar{\mu}_k) \\ &\quad - P(A_2)\bar{\mu}_i - P(A_2)\bar{\mu}_k \\ &= P(S_i)\bar{\mu}_i + P(S_k)\bar{\mu}_k - P(S_i \cap S_k)(\bar{\mu}_i + \bar{\mu}_k - \bar{\mu}_i \cap \bar{\mu}_k) \\ &= P(S_i)\bar{\mu}_i + P(S_k)\bar{\mu}_k - P(S_i \cap S_k)(\bar{\mu}_i \cup \bar{\mu}_k) \\ &= f(S_i) + f(S_k) - f(S_i \cap S_k) \end{aligned} \quad (4.13)$$

$$= P(\bar{C}_i)\bar{\mu}_i + P(\bar{C}_k)\bar{\mu}_k - P(\bar{C}_i \cap \bar{C}_k)(\bar{\mu}_i \cup \bar{\mu}_k) \quad (4.14)$$

In the above expression $\bar{\mu}_i \cup \bar{\mu}_k$ is the summation of the repair rates of the components which belong either to C_i or to C_k or to both i.e.

$$\bar{\mu}_i \cup \bar{\mu}_k = \sum_{j \in C_i \cup C_k} \mu_j$$

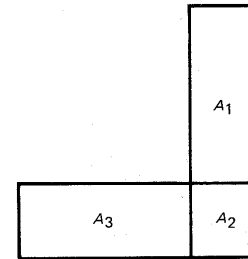


Fig. 4.11 Venn diagram for overlapping subsets representing two minimal cut sets

Equations (4.13) and (4.14) can be extended to the union of three subsets representing three minimal cut sets as follows

$$\begin{aligned} f(S_i \cup S_k \cup S_l) &= f(S_i \cup S_k) + f(S_l) - f((S_i \cup S_k) \cap S_l) \\ &= f(S_i) + f(S_k) - f(S_i \cap S_k) + f(S_l) \\ &\quad - f((S_i \cap S_l) \cup (S_k \cap S_l)) \end{aligned}$$

$$\begin{aligned}
&= f(S_i) + f(S_k) + f(S_l) - f(S_i \cap S_k) \\
&\quad - [f(S_i \cap S_l) + f(S_k \cap S_l) - f(S_i \cap S_k \cap S_l)] \\
&= f(S_i) + f(S_k) + f(S_l) \\
&\quad - [f(S_i \cap S_k) + f(S_i \cap S_l) + f(S_k \cap S_l)] \\
&\quad + f(S_i \cap S_k \cap S_l) \\
&= P(\bar{C}_i)\bar{\mu}_i + P(\bar{C}_k)\bar{\mu}_k + P(\bar{C}_l)\bar{\mu}_l \\
&\quad - [P(\bar{C}_i \cap \bar{C}_k)(\bar{\mu}_i \cup \bar{\mu}_k) + P(\bar{C}_k \cap \bar{C}_l)(\bar{\mu}_k \cup \bar{\mu}_l) \\
&\quad + P(\bar{C}_i \cap \bar{C}_l)(\bar{\mu}_i \cup \bar{\mu}_l)] \\
&\quad + P(\bar{C}_i \cap \bar{C}_k \cap \bar{C}_l)(\bar{\mu}_i \cup \bar{\mu}_k \cup \bar{\mu}_l)
\end{aligned}$$

where as before $\bar{\mu}_i \cup \bar{\mu}_k \cup \bar{\mu}_l$ is the summation of the repair rates of components which belong to any or all of the minimal cut sets C_i, C_k and C_l . Proceeding in the above manner, the formula can be extended to the general case of m cut sets

$$\begin{aligned}
f_f &= f(S_1 \cup S_2 \cup S_3 \cup \dots \cup S_m) \\
&= [P(\bar{C}_1)\bar{\mu}_1 + P(\bar{C}_2)\bar{\mu}_2 + \dots + P(\bar{C}_m)\bar{\mu}_m] \leftarrow \binom{m}{1} \text{ terms order 1} \\
&\quad - [P(\bar{C}_1 \cap \bar{C}_2)(\bar{\mu}_1 \cup \bar{\mu}_2) + P(\bar{C}_1 \cap \bar{C}_3)(\bar{\mu}_1 \cup \bar{\mu}_3) + \dots \\
&\quad + P(\bar{C}_i \cap \bar{C}_j)(\bar{\mu}_i \cup \bar{\mu}_j)] \leftarrow \binom{m}{2} \text{ terms order 2} \\
&\quad + [P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3)(\bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_3) + \dots \\
&\quad + P(\bar{C}_i \cap \bar{C}_j \cap \bar{C}_k)(\bar{\mu}_i \cup \bar{\mu}_j \cup \bar{\mu}_k)] \leftarrow \binom{m}{3} \text{ terms order 3} \\
&\quad \dots \dots \dots \\
&\quad (-1)^{m-1} [P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3 \cap \dots \cap \bar{C}_m) \\
&\quad \times (\bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_3 \cup \dots \cup \bar{\mu}_m)] \quad \text{order } m
\end{aligned} \tag{4.15}$$

The calculation of the terms of the above expansion involves simple multiplications and additions. The contribution by terms beyond the third or

fourth order is quite insignificant and the calculations can be suitably truncated. In high reliability systems the first row alone gives good results

i.e.

$$f_{U1} \simeq \sum_i P(\bar{C}_i)\bar{\mu}_i$$

This equation gives the upper bound approximation for the frequency of system failure. The lower bound to the frequency is

$$f_{L1} = f_{U1} - \sum_{i < k} P(\bar{C}_i \cap \bar{C}_k)(\bar{\mu}_i + \bar{\mu}_k).$$

f_{U1} and f_{L1} are the first upper and lower bounds to the approximation of f_f given by Equation (4.15). Successively closer alternating bounds can be obtained by the additions of odd and even order terms. The computation can be truncated when the margin between upper and lower bounds becomes negligible.

The use of the cut set manipulation equations can be illustrated by application to the network of Fig. 4.8. The minimal cuts for it have already been enumerated in Table 4.1. All the components are assumed identical each having λ and μ as the failure and repair rates. The probability of component failure

$$P_d = \frac{\lambda}{\lambda + \mu}$$

The expression for probability of system failure is

$$\begin{aligned}
P_f &= P(\bar{C}_1 \cup \bar{C}_2 \cup \bar{C}_3 \cup \bar{C}_4) \\
&= [P(\bar{C}_1) + P(\bar{C}_2) + P(\bar{C}_3) + P(\bar{C}_4)] \\
&\quad - [P(\bar{C}_1 \cap \bar{C}_2) + P(\bar{C}_1 \cap \bar{C}_3) + P(\bar{C}_1 \cap \bar{C}_4) + P(\bar{C}_2 \cap \bar{C}_3) \\
&\quad + P(\bar{C}_3 \cap \bar{C}_4) + P(\bar{C}_2 \cap \bar{C}_4)] \\
&\quad + [P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3) + P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_4) + P(\bar{C}_2 \cap \bar{C}_3 \cap \bar{C}_4) \\
&\quad + P(\bar{C}_1 \cap \bar{C}_3 \cap \bar{C}_4)] - [P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3 \cap \bar{C}_4)]
\end{aligned}$$

The various terms are evaluated below

$$\begin{aligned}
P(\bar{C}_1) &= P(\bar{C}_2) = (p_d)^2 \\
P(\bar{C}_3) &= P(\bar{C}_4) = (p_d)^3 \\
P(\bar{C}_1 \cap \bar{C}_2) &= P(\bar{C}_1 \cap \bar{C}_3) = P(\bar{C}_1 \cap \bar{C}_4) = P(\bar{C}_2 \cap \bar{C}_3) \\
&= P(\bar{C}_2 \cap \bar{C}_4) = (p_d)^4
\end{aligned}$$

$$\begin{aligned}
P(\bar{C}_3 \cap \bar{C}_4) &= (p_d)^5 \\
P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3) &= P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_4) = P(\bar{C}_2 \cap \bar{C}_3 \cap \bar{C}_4) \\
&= P(\bar{C}_1 \cap \bar{C}_3 \cap \bar{C}_4) = (p_d)^5 \\
P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3 \cap \bar{C}_4) &= (p_d)^5
\end{aligned}$$

Substituting these values

$$P_f = 2p_d^2 + 2p_d^3 - 5p_d^4 + 2p_d^5$$

Calculation of frequency of failure.

The formula for failure frequency can be written as

$$\begin{aligned}
f_f &= [P(\bar{C}_1)\bar{\mu}_1 + P(\bar{C}_2)\bar{\mu}_2 + P(\bar{C}_3)\bar{\mu}_3 + P(\bar{C}_4)\bar{\mu}_4] \\
&\quad - [P(\bar{C}_1 \cap \bar{C}_2)(\bar{\mu}_1 \cup \bar{\mu}_2) + P(\bar{C}_1 \cap \bar{C}_3)(\bar{\mu}_1 \cup \bar{\mu}_3) \\
&\quad + P(\bar{C}_1 \cap \bar{C}_4)(\bar{\mu}_1 \cup \bar{\mu}_4) + P(\bar{C}_2 \cap \bar{C}_3)(\bar{\mu}_2 \cup \bar{\mu}_3) \\
&\quad + P(\bar{C}_3 \cap \bar{C}_4)(\bar{\mu}_3 \cup \bar{\mu}_4) + P(\bar{C}_2 \cap \bar{C}_4)(\bar{\mu}_2 \cup \bar{\mu}_4)] \\
&\quad + [P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3)(\bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_3) \\
&\quad + P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_4)(\bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_4) \\
&\quad + P(\bar{C}_2 \cap \bar{C}_3 \cap \bar{C}_4)(\bar{\mu}_2 \cup \bar{\mu}_3 \cup \bar{\mu}_4) \\
&\quad + P(\bar{C}_1 \cap \bar{C}_3 \cap \bar{C}_4)(\bar{\mu}_1 \cup \bar{\mu}_3 \cup \bar{\mu}_4)] \\
&\quad - [P(\bar{C}_1 \cap \bar{C}_2 \cap \bar{C}_3 \cap \bar{C}_4)] (\bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_3 \cup \bar{\mu}_4)
\end{aligned}$$

Now

$$\bar{\mu}_1 = \bar{\mu}_2 = 2\mu$$

$$\bar{\mu}_3 = \bar{\mu}_4 = 3\mu$$

$$\bar{\mu}_1 \cup \bar{\mu}_2 = \bar{\mu}_1 \cup \bar{\mu}_3 = \bar{\mu}_1 \cup \bar{\mu}_4 = \bar{\mu}_2 \cup \bar{\mu}_3 = \bar{\mu}_2 \cup \bar{\mu}_4 = 4\mu$$

$$\bar{\mu}_3 \cup \bar{\mu}_4 = 5\mu$$

$$\bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_3 = \bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_4 = \bar{\mu}_2 \cup \bar{\mu}_3 \cup \bar{\mu}_4 = \bar{\mu}_1 \cup \bar{\mu}_3 \cup \bar{\mu}_4 = 5\mu$$

and

$$\bar{\mu}_1 \cup \bar{\mu}_2 \cup \bar{\mu}_3 \cup \bar{\mu}_4 = 5\mu$$

Substituting these values

$$f_f = (4p_d^2 + 6p_d^3 - 20p_d^4 + 10p_d^5)\mu$$

The numerical values were obtained by keeping μ equal to 438 repairs per year, i.e. a mean down time of the component equal to 20 hours, and varying the failure rate from two failures per year to 219 failures per year. These results are shown in Table 4.2. The purpose of this study is to show the difference between exact and upper bound approximate results as a function of the component reliability. It can be seen that when component reliability is close to unity, the upper bound approximations gives almost exact results.

Algorithm to Determine Minimal Cut Sets

When the reliability block diagram is small, the set of minimal cuts can be found by visual examination. In a large network such an examination could be very laborious and perhaps even impossible. Reference 3 describes an algorithm for generating the minimal cuts of a network. This procedure is quite suitable for computer application.

The input or supply node of a reliability block diagram will be denoted by s and the output or the load node by l . The removal of the components in a minimal cut separates the network into exactly two connected subnetworks, one containing the node s and the other node l . The minimal cut in other words partitions the set N of all nodes into subsets N^+ and N^- . The set N^+ defines a connected subnetwork that includes s and N^- defines another connected subnetwork that includes l .

The algorithm generates a tree of the network, the vertices being the points and the edges being the line segments on the tree. The tree starts from the root vertex denoted by 0. The edges are marked n^+ or n^- . The symbol n^+ means that the node n of the reliability diagram is a member of N^+ and n^- means that node n of reliability diagram is a member of N^- . With each node i of the tree are associated four subsets of the nodes of the reliability diagram, X_{1i} , X_{2i} , X_{3i} and W_i . Let h_i be the unique simple path connecting vertex i to the root vertex, then

Node $n \in X_{1i}$ if an edge in the path h_i is labelled n^+

Node $n \in X_{2i}$ if an edge in the path h_i is labelled n^-

Node $n \in X_{3i}$ if it is neither in X_{1i} nor in X_{2i} but is a member of N .

Node $n \in W_i$ if it is in X_{3i} and if it is the terminal of a component whose other terminal is a member of X_{1i}

Table 4.2 Probabilities and frequencies of failure for the system in Fig. 4.8

 $\mu = 438.0$ repairs/year

λ failures per year	Component Reliability	Probability of Failure		Frequency of Failure	
		Exact	Under Bound Approximately	Exact	Under Bound Approximately
2-0	0.995455	$4150801221 \times 10^{-4}$	$4151014274 \times 10^{-4}$	$3644142301 \times 10^{-1}$	364451540×10^{-1}
4-0	0.990950	$1652457187 \times 10^{-3}$	$1652791341 \times 10^{-3}$	1453752889	145433779
8-0	0.982063	$6545162018 \times 10^{-3}$	$6550300836 \times 10^{-3}$	5779632215	578861915
16-0	0.964758	$2563976981 \times 10^{-2}$	$2571581307 \times 10^{-2}$	2277773988×10^1	22910491×10^1
32-0	0.9319149	$9797874590 \times 10^{-2}$	$9902391570 \times 10^{-2}$	8769140832×10^1	895097294×10^1
219-0	0.6666667	2427983539	2962962962	2018765432×10^3	2920000000×10^3

The algorithm now proceeds as follows:

Step 1: Generate vertices 0, 1 and 2 and label the edges (0, 1) and (1, 2) as s^+ and 1^- respectively. This means that for all minimal cuts s and 1 are members of N^+ and N^- respectively. The vertices 0 and 1 are assumed scanned but vertex 2 is unscanned. Go to step 2.

Step 2: If there are no unscanned vertices, the complete tree has been generated and the algorithm terminates. Otherwise, choose the unscanned vertex with the greatest index and mark it scanned and let it be called i . Find the unique simple path h_i from i to the root vertex 0, and identify the subsets X_{1i} , X_{2i} , X_{3i} and W_i as defined above. If W_i is null, i.e. has no member go to step 7, otherwise choose x , an element of W_i and construct the subnetwork defined by the set of nodes $X_4 = X_{2i} \cup X_{3i} - x$. If this subnetwork is connected, go to step 3, and if not go to step 4.

Step 3: Generate two new vertices k and $k+1$ where k is one greater than the highest index so far in the tree. Create edges (i, k) and $(i, k+1)$ and label them x^+ and x^- . The vertices k and $k+1$ are unscanned vertices. Go to step 2.

Step 4: Find the set of nodes X_5 which defines a connected subnetwork that includes 1. If X_{2i} is a subset of X_5 go to step 5, otherwise to step 6.

Step 5: Generate vertex k and edge (i, k) labelled x^+ . Determine the set $X_6 = X_4 - X_5$. If $|X_6|$ is the number of the elements of X_6 , then create vertices $k+1, k, \dots, k+|X_6|$ and generate edges $(k, k+1), (k+1, k+2), \dots, (k+|X_6|-1, k+|X_6|)$ and label them Z^+ where $Z \in X_6$. Finally generate vertex $k+|X_6|+1$ and create edge $(i, k+|X_6|+1)$ labelled x^- . Go to step 2.

Step 6: Generate one new vertex index k and create the edge (i, k) labelled x^- . Go to step 2.

Step 7: At this step a minimal cut has been generated. The set $N_i^+ = X_{1i}$ and $N_i^- = N - X_{1i}$. The components whose one terminal is in N_i^+ and the other in N_i^- are the members of the minimal cut. Go to step 2 to create other minimal cuts.

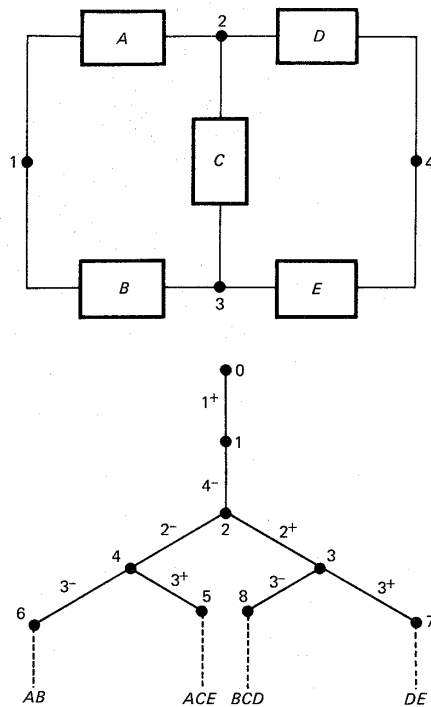
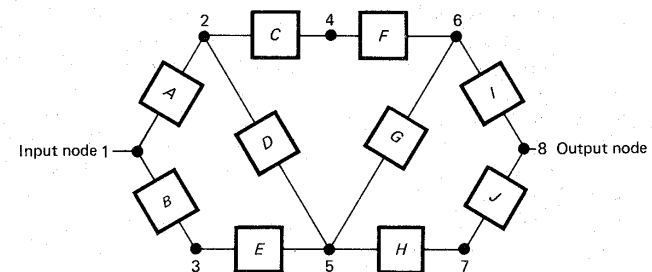


Fig. 4.12 The minimal cut generation tree

The above algorithm is illustrated by application to the reliability block diagram of Fig. 4.8. As the nodes are to be marked by integers the components may be represented by letters as shown in Fig. 4.12 which shows the tree for this network. The reader is urged to verify this tree by going through the different steps of the algorithm.

Exercises

1. Calculate the probability and frequency of system failure in Example 4.2 using the state space approach.
2. Enumerate the minimal cut sets for the following reliability block diagram.



References

1. R. Billinton and C. Singh, Generating Capacity Reliability Evaluation in Interconnected Systems Using a Frequency and Duration Approach—Part I. Mathematical Analysis, *IEEE Trans., Power App. and Systems* (July/August 1971).
2. J.A. Buzacott, Network Approaches to Finding the Reliability of Repairable Systems, *IEEE Trans. on Reliability*, R-19, 4 (November 1970).
3. P.A. Jenson and M. Bellmore, An Algorithm to Determine the Reliability of a Complex Systems, *IEEE Trans. on Reliability*, R-18, 4 (November 1969).
4. C. Singh and R. Billinton, A New Method to Determine the Failure Frequency of a Complex System, *Micro-Electronics and Reliability*, Vol. 12, No. 5 (1973).
5. C. Singh, 'Reliability Modelling and Evaluation in Electric Power Systems', Ph.D. Thesis, University of Saskatchewan, Saskatoon, Canada (August 1972).
6. L.M. Shooman, *Probabilistic Reliability*, New York, McGraw-Hill (1968).

CHAPTER 5

Techniques for Large Systems

Introduction

Several mathematical modelling techniques suitable for reliability evaluation have been illustrated. Although the problem of system reliability may appear conceptually simple, at least when constant transition rates can be assumed, the task may become difficult with large and complex systems. In control systems, communication networks and power networks, it is sometimes easy to derive the reliability block diagrams from the schematic diagrams. When the network approach can be used, the problem becomes relatively simple. In certain cases, however, it is difficult and often impossible to apply this approach. This is especially so with systems involving dependent failure or repair modes and those involving graded modes of operation. In such cases the state space approach is often the only method available. This chapter will identify problem areas while using this approach and suggest suitable techniques for overcoming these difficulties.

The Problem Areas

It was pointed out in the last chapter that the state space approach essentially involves the following steps:

1 *Evolution of the state space and the interstate transition rates.*

In a small system, it is possible to draw the state transition diagram of the system and then solve it with a calculator or program it for the digital computer. When, however, the state space is large, this procedure becomes impractical and often impossible. In such cases this process can be performed by a computer. The method was described in the last chapter. The basic idea is to let the states sequentially evolve by the realization of each possible transition mode of the components. When the transition modes of the components are dependent, system states may impose restrictions on some transition modes but in systems consisting of independent components, each component is allowed to realize all of its transition modes. In certain systems using the symmetry of the transition diagrams, special methods may be evolved for generating the state space and the interstate transition rates. The main problem in this step is to ensure that the correct state space and interstate transition rates have been generated. The generation of a correct state transition rate matrix is very

important as this is the foundation for further calculations.

In some cases, the program may be used to generate the transition rate matrix of a similar but smaller system. The matrix along with the description of states can be printed out and checked visually. For ease in checking, fictitious transition rates, usually whole numbers may be used. When this is not possible or convenient, the program should have an independent checking subroutine. Each system state should be examined for the possible modes of transition. For each mode the resulting state description should be constructed and compared with the ones which have already been generated. If both the state description and the interstate transition rate agree, for all the states, it is reasonable to assume that the state transition rate matrix is correct. If at any state there is any discrepancy, the state and the resulting states should be printed. This will help in debugging.

The other problem is the size of the state space. As an example, a system consisting of n independent two state units will have 2^n possible states. The state transition matrix will be of $2^n \times 2^n$ size. The available memory in the computer may soon be exhausted. This difficulty can be alleviated to some extent by using the principles of sparsity programming. This problem will, however, be discussed in more detail later.

2 *Calculation of the state probabilities*

When the components are independent, the problem becomes simpler, since the system state probabilities can be derived from the component state probabilities by the simple multiplication rule of probabilities. Since the calculation of the probability of each system state can be done independently of the other states, evaluation can be made selectively for the states required for final calculation of the reliability measures. In the case of dependent transition modes the probabilities of all the states have to be calculated since the states can be solved independently for probabilities. If time dependent solutions are required, the state space equation has to be solved. The methods of doing it have already been discussed. Most often, however, the steady state solution is required. The problem is then reduced to solving a set of simultaneous linear equations. This is usually done using the Gauss elimination or Gauss–Jordan method. These techniques are explained in Appendix I. Though conceptually simple, the problem can become formidable even on the computer as the number of components becomes large, and the state space becomes very large. Computer storage limitations, the errors introduced by rounding off and the computation time required all make the problem a very difficult one. It should be appreciated that the transition rate matrix contains failure rates which are usually very small as compared to the repair rates which are comparatively large. The operations on small and large numbers are bound to introduce rounding off errors. It can therefore be appreciated that steps have to be taken to limit the state space.

3 Calculation of the reliability measures

The two key indices are the probability and frequency of encountering a certain configuration of states. The other indices, the mean cycle time and the mean down time can be then simply derived. Success or failure may not provide a complete description and it may be necessary to evaluate reliability measures for graded modes of operation e.g. in a power system it may be necessary to calculate the probabilities and frequencies of having different magnitudes of capacity deficiencies. Similarly in a transit system it may be necessary to know the probabilities and frequencies of having various numbers of vehicles available for transportation. The various states are then grouped into subsets denoting a particular condition. The probabilities and frequencies of these subsets are then calculated using the equations given in Chapter 3. In general all the states have to be scanned to classify them into the required subsets and to select the appropriate transition rates for frequency calculations. In many cases, the classification can be done by taking the advantage of the systematic pattern of the state space. But if all the states have to be scanned, it could be a time-consuming process.

It can be seen from the above discussion that the problem with large systems is essentially the size of the state space. In some situations the size of the state space does not grow proportionately with the number of components. The growth of the state space in such cases is restricted because of the dependency considerations. A familiar example is that of a series system of n components when the assumption is made that the exposure of components to failure is zero when the system is down. In this case the number of states is simply $n+1$ as compared with 2^n when the component failures are independent.

In general, the following procedure should be adopted.

1. The system should be divided into suitable subsystems which can be handled conveniently one at a time. A system is usually naturally subdivided into subsystems on the layout or functional basis. Most often this natural subdivision can be used as the basis for classification for reliability evaluation, but it is not necessary to do so. Primary concern is on the system effect of component failures. Every attempt should be made to divide the system into independent subsystems. The advantage in doing so is that the probabilities of the system can then be found by simple multiplication of the probabilities of the states of the subsystems. Another advantage is that the combination of the independent subsystems is simpler and the equivalent transition rate concept can be more conveniently employed. This will become clearer when the implications of the equivalent transition rate are considered. The independence of various subsystems may sometimes be achieved by shifting some components from one natural subsystem to another. This trick can sometimes prove quite useful.
2. The state space of each subsystem may be reduced either by merging states

or by truncating very low probability states. The principles of both of these techniques are explained in detail in this chapter.

3. The subsystems should then be combined into a complete system and the required reliability measures evaluated.

Two important concepts in large systems are therefore the merging of the states and truncation of states.

Equivalent Transition Rate and the Conditions of Mergeability

The equivalent transition rate concept was introduced in Chapter 3 and used for independent subsystems in Chapter 4, while using the network reduction approach. The principal use of the equivalent transition rate is in reducing the system or subsystem state space. The basic idea is to find a state space which is equivalent to the original state space but is more convenient to use. Assume that the entire state space is partitioned into m subsets

$X_i, i = 1, 2, \dots, m$. The equivalent transition rate from subset X_p to X_q is obtained by using Equation (3.37)

$$\lambda_{pq}^{(e)}(t) = \frac{\sum_{i \in X_p} \sum_{j \in X_q} p_i(t) \lambda_{ij}}{\sum_{i \in X_p} p_i(t)} \quad (5.1)$$

This section examines the conditions under which this equivalent model will give the same results as would be obtained by using the original model. Complete knowledge of these conditions is quite important and the lack of awareness in this area can lead to gross errors. The concept of equivalent transition rate will be examined both in the transient domain and under the equilibrium conditions. This concept can be useful in the following ways:

(a) System State Space Reduction

Sometimes it may be desirable to reduce the state space of the system by merging together certain sets of states. The equivalent transition rates among the subsets of the lumped states are required.

State merging in the system state space is generally of value when the equivalent transition rates can be calculated without having to solve for the system state probabilities.

(b) Subsystem State Space Reduction

The most practical method is to break down the system into subsystems which can be individually solved and then to combine these solutions to get the results for the entire system. It may be desirable to reduce the state space of an individual subsystem and therefore the equivalent transfer rate will be calculated.

Model reduction in this case is useful even if all of the subsystem state probabilities have to be calculated to determine the equivalent transfer rates.

Transient Domain

The equivalent transfer rate is first examined from the point of view of system state space reduction. The equivalent transition rate from subset X_p to X_q is given by Equation (5.1).

The following observations can be made from the relationship in (5.1):

1. Since the state probabilities in Equation (5.1) depend upon the initial state probability vector, the equivalent transfer rate, $\lambda_{pq}^{(e)}(t)$ is a function of the initial state probability vector. In contrast to this the interstate transition rates λ_{ij} of the original state space are independent of any such condition. If $\lambda_{pq}^{(e)}(t)$ are to be independent of such a restriction they must then be independent of the state probabilities. This can happen if $\sum_{j \in X_q} \lambda_{ij}$ is the same for all $i \in X_p$ so that

$$\begin{aligned} \lambda_{pq}^{(e)}(t) &= \frac{\sum_{j \in X_q} \lambda_{ij} \sum_{i \in X_p} p_i(t)}{\sum_{i \in X_p} p_i(t)} \\ &= \sum_{j \in X_q} \lambda_{ij} = \lambda_{pq}^{(e)} \end{aligned} \quad (5.2)$$

2. Unless the expression for equivalent transfer rate is independent of the state probabilities as in Equation (5.2), $\lambda_{pq}^{(e)}(t)$ is obviously a function of t . If an explicit expression for $\lambda_{pq}^{(e)}(t)$ can be obtained, the solution for the reduced state space can be obtained by using Kolmogorov differential equations for the time specific transition rate matrix

$$A(u)P(t, u) = \frac{\partial P(t, u)}{\partial u} \quad (5.3)$$

where for $u > t$

$P(t, u)$ = The column matrix whose i th value $p_i(t, u)$ represents the probability of being in state i at time u for the given initial condition at t , i.e.

$$p_i(t, u) = P\{Z(u) = i | Z(t) = j\}$$

and $A(u)$ = The transpose of the time specific transition rate matrix. When the initial conditions are specified at the time of origin

$$A(u)P(0, u) = \frac{\partial P(0, u)}{\partial u} \quad (5.4)$$

When Equation (5.4) is used for the reduced system, the initial condition, $Z(0)$ should be the same as for the $\lambda_{pq}^{(e)}(u)$ in (5.1)

Writing the differential equation for the lumped state q

$$\frac{\partial p_q(0, u)}{\partial u} = \sum_{p \in X_r} \lambda_{pq}^{(e)}(u) p_p(0, u) - p_q(0, u) \sum_{p \in X_r} \lambda_{qp}^{(e)}(u) \quad (5.5)$$

where $X_r = \{\text{All states except } q\}$, the subscript r indicating that the reference is to the reduced state space.

Since the initial condition for (5.1) and (5.4) is the same

$$p_p(0, u) = \sum_{i \in X_p} p_i(u)$$

and

$$\frac{\partial p_q(0, u)}{\partial u} = \sum_{j \in X_q} p_j'(u)$$

The indices p and q refer to the lumped states whereas i and j indicate the original states. Substituting these values into Equation (5.5) and substituting the values of $\lambda_{pq}^{(e)}(t)$ from Expression (5.1)

$$\sum_{j \in X_q} p_j'(u) = \sum_{p \in X_r} \sum_{i \in X_p} \sum_{j \in X_q} p_i(u) \lambda_{ij} - \sum_{p \in X_r} \sum_{j \in X_q} \sum_{i \in X_p} p_j(u) \lambda_{ji}$$

Now if $X^+ = X_q$ and X^- is the disjoint subset, then the above equation can be written as

$$\sum_{j \in X^+} p_j'(u) = \sum_{i \in X^-} \sum_{j \in X^+} p_i(u) \lambda_{ij} - \sum_{j \in X^+} \sum_{i \in X^-} p_j(u) \lambda_{ji}$$

It can be seen that this is the same expression as would be obtained by following the argument used to derive Equation (3.7). The time specific equivalent transfer rates, therefore, do represent the process accurately but the essential point is that this does not solve anything because all the state probabilities have to be found before the equivalent transfer rates can be determined. The only case when the state probabilities do not have to be evaluated is the one given by Equation (5.2). In summary, mergeability from a systems state space viewpoint can be defined as follows:

Definition: If the entire state space is partitioned into m subsets X_i , $i = 1, 2, \dots, m$ and the equivalent transfer rate from subset X_p to X_q , given by Equation (5.1), is time invariant and independent of the initial state probability vector, then the state space is said to be mergeable into the said partition.

The necessary and sufficient condition is that the transition rate from each state in subset X_p to each of the states in X_q when summed over all the states in X_q is the same for each state in X_p and the required equivalent rate is given by

$$\lambda_{pq}^{(e)} = \sum_{j \in X_q} \lambda_{ij}$$

When examining mergeability from the point of view of subsystem reduction, the state space of a subsystem S_a is assumed to be partitioned into subsets $X_1^a, X_2^a, \dots, X_p^a$ in such a manner that no information about the reliability analysis is lost. The equivalent transfer rate from one subset to another can be obtained using Equation (5.1). The important consideration for the present purpose, however, is that this equivalent rate should hold when this subsystem is combined with another subsystem. The equivalent transfer rate from the lumped states of subset X_l^a to those of X_m^a is given by

$$\lambda_{lm}^a(t) = \left(\sum_{i \in X_l^a} p_i(t) \sum_{j \in X_m^a} \lambda_{ij} \right) / \sum_{i \in X_l^a} p_i(t)$$

After merging together the states in each subset, there will be p equivalent states in the subsystem S_a , one equivalent state corresponding to one subset. This subsystem is now combined with another subsystem S_b with equivalent states $b_1, b_2, \dots, b_q$. It is assumed that in the combined system there will be $p \times q$ states, there being a whole set of equivalent states a_u , $u = 1, \dots, p$, corresponding to every equivalent state in subsystem b . This can be indicated as follows

$$\begin{array}{cccc} (b_1)a_1, & (b_1)a_2, & \dots, & (b_1)a_p \\ (b_2)a_1, & (b_2)a_2, & \dots, & (b_2)a_p \\ \vdots & \vdots & \ddots & \vdots \\ (b_q)a_1, & (b_q)a_2, & \dots, & (b_q)a_p \end{array}$$

It should be appreciated that this arrangement may not always be possible as some combinations may be incompatible. In the subsystems exposed to a

common two state fluctuating environment, which may alter the component failure and repair rates, such an arrangement will exist in either state. Only the arrangement shown above will be discussed but the results will also hold for other cases. After combination of the two subsystems, certain interstate transition rates may be altered but it is assumed that the interstate transition rates with respect to the merged states are not changed so that the act of combining the subsystems does not affect the information contained in the merged states.

For the transfer rates to hold after combining, the equivalent transfer rate from $(b_u)a_l$ to $(b_u)a_m$ should be the same as that from a_l to a_m . Here u may be any number from 1 to q . Therefore, for mergeability

$$\lambda_{lm}^{ab}(t) = \lambda_{lm}^a(t)$$

i.e.,

$$\frac{\sum_{i \in X_l^a} p_i^u(t) \sum_{j \in X_m^a} \lambda_{ij}}{\sum_{i \in X_l^a} p_i^u(t)} = \frac{\sum_{i \in X_l^a} p_i(t) \sum_{j \in X_m^a} \lambda_{ij}}{\sum_{i \in X_l^a} p_i(t)} \quad (5.6)$$

where

$p_i^u(t)$ = The probability, for the given initial state, of being in state $i \in X_l^a$ corresponding to the condition represented by b_u in subsystem S_b .

The following conclusions can be drawn from the Equality (5.6):

(1) For independent subsystems

$$p_i^u(t) = p_i(t) p_{bu}(t)$$

where

$p_{bu}(t)$ = The probability of being in the equivalent state b_u for the given initial condition

Therefore

$$\begin{aligned} \lambda_{lm}^{ab}(t) &= \left(\sum_{i \in X_l^a} p_i(t) p_{bu}(t) \sum_{j \in X_m^a} \lambda_{ij} \right) / \sum_{i \in X_l^a} p_i(t) p_{bu}(t) \\ &= \left(\sum_{i \in X_l^a} p_i(t) \sum_{j \in X_m^a} \lambda_{ij} \right) / \sum_{i \in X_l^a} p_i(t) \\ &= \lambda_{lm}^a(t) \end{aligned}$$

The time specific equivalent transfer rate, therefore, holds unconditionally if the subsystems being combined are independent. The evaluation of these rates is, of course, a separate problem.

(2) If, however, the subsystems are not independent, the Equality (5.7) can hold if

(i) $\sum_{j \in X_m^a} \lambda_{ij}$ is the same for any i in which case

$$\begin{aligned} \lambda_{im}^{ab}(t) &= \left(\sum_{j \in X_m^a} \lambda_{ij} \right) \frac{\sum_{i \in X_l^a} p_i^u(t)}{\sum_{i \in X_l^a} p_i^u(t)} \\ &= \sum_{j \in X_m^a} \lambda_{ij} \\ &= \lambda_{im}^a \end{aligned}$$

This is the same condition as previously specified for mergeability.

(ii) Alternatively if for a given initial state probability vector

$$\begin{aligned} p_1^u(t) &= p_2^u(t) = \dots = p_i^u(t) \quad \text{which implies} \\ p_1(t) &= p_2(t) = \dots = p_i(t) \end{aligned}$$

i.e., if the states being lumped have the same time specific probability

$$\begin{aligned} \lambda_{im}^{ab}(t) &= \lambda_{im}^a \\ &= \sum_{i \in X_l^a} \sum_{j \in X_m^a} \lambda_{ij} / n_l \end{aligned}$$

where n_l = The number of states in subset X_l^a

Steady state is a special case of transient analysis with $t \rightarrow \infty^+$. The analysis for steady state conditions of mergeability is the same as that for the transient case with the following differences:

1. The steady state equivalent transfer rate is time invariant.
2. The steady state probabilities are independent of the initial conditions and therefore the steady state equivalent transfer rate is independent of such a restriction.

The results of the mergeability analysis are summarized below.

1. If $\sum_{j \in X_m^a} \lambda_{ij}$ is the same for any $i \in X_l^a$, then the equivalent transfer rate is time invariant and independent of the initial condition. When this condition is satisfied between all the subsets taken in pairs, the Markov process of the

subsystem S_a is said to be mergeable into these disjoint subsets. The equivalent transfer rates so determined hold when this subsystem is combined with another subsystem S_b , provided the interstate transition information contained in the lumped states does not change by virtue of this combination.

2. If for the given initial condition, the probabilities of the states in a subset are equal, then these states are mergeable and the equivalent transition rate holds when this subsystem is combined with another subsystem, provided the initial condition is not violated. In the steady state case, the initial condition does not affect the conclusion.

3. For independent subsystems, the states of the subsystem may be lumped into any desired disjoint subsets. The equivalent transition rate is unaltered by combination but because of the computation effort required, this does not seem to be of much significance in the time specific analysis. In steady state analysis, this can facilitate the calculation of the frequency index. A familiar example is the lumping together of the identical capacity outage states in a generation system model before combining it with the load model.

Example 5.1: The concepts outlined in the previous section can be illustrated for the steady state condition using the simple system shown in Fig. 5.1.

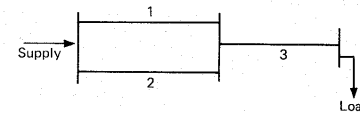


Fig. 5.1 A simple three unit series-parallel system

The system consists of two lines 1 and 2 in parallel. This combination is then in series with component 3. Both lines 1 and 2 have identical failure and repair rates λ and μ and each is capable of supplying the full load. The failure and repair rates of line 3 are λ_3 and μ_3 respectively. The system is now assumed to consist of subsystem S_a of lines 1 and 2 and the subsystem S_b of line 3. The state transition diagrams of the two subsystems taken individually are shown in Figs. 5.2 (a) and (b) respectively. On combining S_a and S_b , there will be a total of

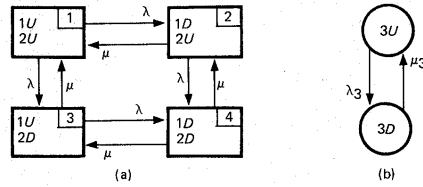


Fig. 5.2 (a) The state transition diagram of Subsystem S_a
(b) The state transition diagram of Subsystem S_b

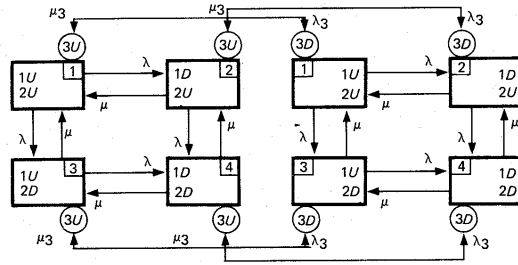


Fig. 5.3 The state transition diagram after combining S_a and S_b

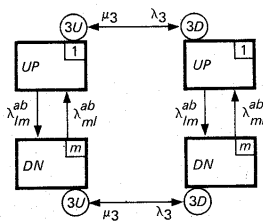


Fig. 5.4 The state transition diagram after combining reduced S_a with S_b

eight states. These can be written as

$$(3U)1, (3U)2, (3U)3, (3U)4$$

$$(3D)1, (3D)2, (3D)3, (3D)4$$

The condition inside the parenthesis is that of S_b and the outside numbers refer to the state numbers of S_a . Only one line out of $\{1, 2\}$ is required for successful operation and therefore the behaviour of S_a can be represented by a two state component in which state l corresponds to states $\{1, 2, 3\}$ and state m to state $\{4\}$ of the original state space. The equivalent transition rates are

$$\lambda_{lm} = \frac{\lambda(p_2 + p_3)}{p_1 + p_2 + p_3}$$

and

$$\lambda_{ml} = 2\mu$$

The question to be examined is, do these equivalent transfer rates hold after combining S_a and S_b ? This problem is considered for both complete and restricted repair facilities.

1 Complete Repair Facilities

When each component can be repaired independently, the two subsystems are independent. It can also be observed that when the two subsystems are independent, the interstate transition modes remain unchanged after interaction. Now, if merging is to be valid, the equivalent transition rate λ_{lm}^{ab} , i.e. from $(3U)l$ to $(3U)m$ should be the same as λ_{lm} , i.e. from l to m . Examining the state transition diagram in Fig. 5.3

$$\lambda_{lm}^{ab} = \frac{(p_2^u + p_3^u)\lambda}{p_1^u + p_2^u + p_3^u}$$

Here the superscript refers to the condition of system b . Since the subsystems are independent

$$\begin{aligned} \lambda_{lm}^{ab} &= \frac{(p_2 p_{3u} + p_3 p_{3u})\lambda}{p_1 p_{3u} + p_2 p_{3u} + p_3 p_{3u}} = \frac{(p_2 + p_3)\lambda}{p_1 + p_2 + p_3} \\ &= \lambda_{lm} \end{aligned}$$

It can be concluded from the above discussion that when S_a and S_b are independent, subsystem S_a can be represented by a two state component having the equivalent transition rates determined by Equation (5.1). This is the basis of the network reduction technique described in the previous chapter. The reason for assuming the subsystems to be independent can now be more fully understood.

2 Restricted Repair Facilities

Now suppose that only two lines can be repaired at a time, then in the event of failure of all the three lines, state (3D)4, in Fig. 5.3, the repair of one line out of $\{1, 2\}$ must wait for the first repair. Referring to Fig. 5.4, λ_{ml}^{ab} from (3U)m to (3U)l is still 2μ but λ_{ml}^{ab} from (3D)m to (3D)l is now μ . The systems are no longer independent and therefore $p_i^u \neq p_i^l, p_i^u$, and consequently $\lambda_{lm}^{ab} \neq \lambda_{lm}$. Merging into the above groupings is not possible.

If, however, states $\{2, 3\}$ are merged to give the equivalent state l and if $\{1\}$ and $\{4\}$ are denoted by n and m respectively, the conditions of mergeability are satisfied and it can be seen that

$$\begin{aligned} \lambda_{ln}^{ab} &= \lambda_{ln} = \mu \\ \text{and} \quad \lambda_{lm}^{ab} &= \lambda_{lm} = \lambda \end{aligned}$$

Merging into these states is therefore still possible.

Components Subject to Fluctuating Environment

The following considers the application of conditions of mergeability to a system consisting of components exposed to a two state fluctuating environment. The two states are designated as N and S states and the durations of these states are assumed to be exponentially distributed. The component failure and repair rates are constant but depend on the environment. Failure and repair of the components are independent in a given environment but the exposure to the common environment introduces the element of interdependence and the probabilities cannot be found by a simple product rule. The conditions of mergeability developed in this chapter will now be applied to this system.

The system is assumed to be divided into a number of subsystems out of which only subsystems S_a and S_b are considered for the sake of convenience. These subsystems are assumed to contain n_a and n_b number of components respectively. Considering S_a , it can exist in 2^{n_a} number of states in each of the environmental states. The component configuration of each state in either of the two environment states is the same but the interstate transition rates in the two weather states are different. The transition rate from a system state in the N environment condition to the corresponding system state in the S environment condition is taken as n and in the reverse direction as m . The states in the N environment may be grouped into subsets $X_1^a, X_2^a, \dots, X_p^a$ and those in the S environment as $X_1^{as}, X_2^{as}, \dots, X_p^{as}$. The combination of states in X_i^a and X_i^{as} is the same. Since the reduced state space of S_a is to be combined with the reduced state space of S_b which is also exposed to the same fluctuating environment, the states in the N and S environment cannot be lumped together. The equivalent

transfer rates can be found by the application of Equation (5.1). It is obvious that the transition rate from X_i^a to X_i^{as} is n and that from X_i^{as} to X_i^a is m . The equivalent transfer rate from X_i^a to X_m^a is given by

$$\lambda_{im}^a = \left(\sum_{i \in X_l^a} p_i \sum_{j \in X_m^a} \lambda_{ij} \right) / \sum_{i \in X_l^a} p_i$$

where

p_i = The probability of being in the i th state in N environment

and

λ_{im}^a = The equivalent transfer rate from the lumped states of subsets X_l^a to those of X_m^a .

The same treatment holds for the states in the S environment. After merging together the states in each subset, there will be p equivalent states in each environment condition in the subsystem S_a , one equivalent state corresponding to one subset. These equivalent states may be indicated by $a_1, a_2, \dots, a_p$ in the N environment condition, and $a_1^s, a_2^s, \dots, a_p^s$ in the S environment condition. Now suppose that this subsystem is combined with the other subsystem S_b with equivalent states $\{b_1, b_2, \dots, b_q\}$ and $\{b_1^s, b_2^s, \dots, b_q^s\}$ in the N and S environment condition. In the combined system there will be $p \times q$ states in each environment condition. This combination for the N environment is shown below.

$$\begin{array}{cccc} (b_1)a_1, & (b_1)a_2, & \dots, & (b_1)a_p \\ (b_2)a_1, & (b_2)a_2, & \dots, & (b_2)a_p \\ \vdots & \vdots & \ddots & \vdots \\ (b_q)a_1, & (b_q)a_2, & \dots, & (b_q)a_p \end{array}$$

Since the two subsystems are exposed to the same fluctuating environment, it can be easily seen that the equivalent transfer rate from $(b_i)a_j$ to $(b_i^s)a_j^s$ is n and that from $(b_i^s)a_j^s$ to $(b_i)a_j$ is m . For steady state mergeability, the transfer rate from $(b_u)a_i$ to $(b_u)a_m$ should be the same as that from a_i to a_m , i.e.,

$$\lambda_{im}^{ab} = \lambda_{im}^a$$

By the applications of the mergeability conditions the following conclusions can be drawn:

1. If S_a and S_b are independent, the equivalent transfer rates hold unconditionally, i.e. the states of the subsystem may be lumped into any desired disjoint subsets. When S_a and S_b are exposed to the same fluctuating environment, the independence is possible if either S_a or S_b or both subsystems have components whose failure and repair rates are the same in both environment

states, i.e. these rates are environment independent.

2. In subsystems which are not independent, the states can be lumped together if, and only if, either

- (i) All the states being merged have the same availability, or
- (ii) The sum of transfer rates from state $i \in X^a$ to all states $j \in X_m^a$, i.e., $\sum_{j \in X_m^a} \lambda_{ij}$ is the same for all i .

The conditions of steady state mergeability restrict the scope of model reduction when the subsystems are not independent but even in this restricted sense model reduction can be of considerable help in large sized networks.

One obvious application is when the elements of a subsystem have identical failure and repair rates. An n identical element subsystem can be lumped into $2(n+1)$ states. For example, a four element subsystem will have the following groups of identical states.

Group number	Number of elements failed	Number of identical states in group
1	0	$\binom{4}{0} = 1$
2	1	$\binom{4}{1} = 4$
3	2	$\binom{4}{2} = 6$
4	3	$\binom{4}{3} = 4$
5	4	$\binom{4}{4} = 1$
Total		16

The 16 states for each environment state can be lumped into five equivalent states. Therefore, the subsystem having 32 states can be adequately represented by 10 states. This application to identical elements subsystems becomes important when it is realized that parallel facilities normally have identical failure and repair parameters.

The merging technique can be illustrated by application to the simple system shown in Fig. 5.5. The data for this system is given below. The N environment corresponds to normal weather and the S environment to stormy or adverse weather.

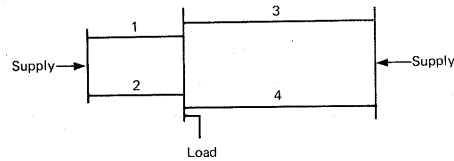


Fig. 5.5 A simple transmission network Mean duration $N = 200$ hours
Mean duration $S = 1.5$ hours

Percentage of failures during S environment = 20%

Components	Average failure rate per year	Mean down time hours
1,2	0.5	5.0
3,4	1.0	10.0

This system can be split into two subsystems A and B having $\{1,2\}$ and $\{3,4\}$ components. The state transition diagram of system i ($= a$ or b) is shown in Fig. 5.6.

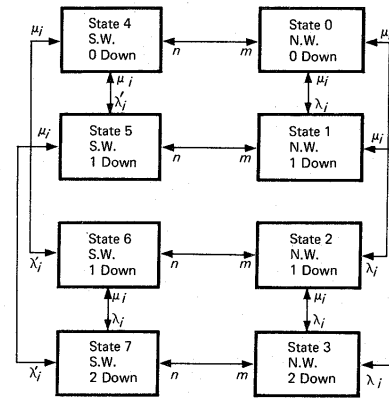


Fig. 5.6: Two identical unit state transition diagram for a two state fluctuating environment.

The following notation is used

$m = \frac{1}{S}$ = Transition rate from Stormy state (SW) to Normal state (NW).

$n = \frac{1}{N}$ = Transition rate from Normal state to Stormy state.

μ_i = The mean repair rate of a component in the i th system. The same repair rate is assumed in normal and stormy states.

λ_i = Normal state failure rate of a component in the i th system.

λ'_i = Stormy state failure rate of a component in the i th system.

The average failure rate λ_{av} is given by

$$\lambda_{av} = \lambda \left[\frac{N}{N+S} \right] + \lambda' \left[\frac{S}{N+S} \right] \quad (5.7)$$

If the number of failures in the stormy weather is x percent

$$\frac{\lambda'S}{\lambda N + \lambda'S} = \frac{x}{100} \quad (5.8)$$

The value of λ and λ' can be determined using Equations (5.7) and (5.8).

Subsystem Reduction

In subsystem i , states 1, 2 and 5, 6 are identical and therefore have equal availabilities. Condition 2(i) is thus satisfied and these states can be lumped together. The reduced model of subsystem A is shown in Fig. 5.7. This reduced model can be combined with the reduced model of system B . The resulting state transition diagram is shown in Fig. 5.8. In this diagram, N, W, S, W stand for normal weather and stormy weather states and XI , ($X = A$ or B and $I = 0, 1, 2$) stands for system X with I components down. It can be seen that in the

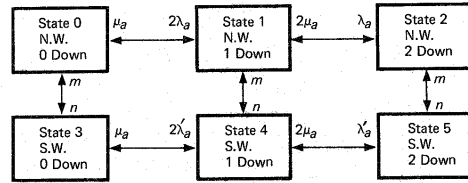


Fig. 5.7 The reduced model of subsystem A

reduced model there are only 18 states as against 32 states in the original model.

The steady state equations can be written in the form

$$TP = B \quad (5.9)$$

where

$$T = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix}$$

where

$$T_{12} = \mu_b [I] \quad (6 \times 6 \text{ matrix})$$

$$T_{13} = 0 \quad (6 \times 6 \text{ matrix})$$

$$T_{21} = \begin{bmatrix} 2\lambda_b [I] & 0 \\ 0 & 2\lambda'_b [I] \end{bmatrix} \quad (6 \times 6 \text{ matrix})$$

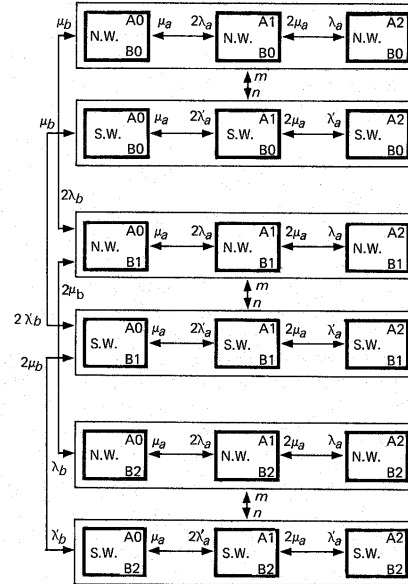


Fig. 5.8 The state transition diagram of the system shown in Fig. 6.1

$$T_{23} = 2 \times T_{12} \text{ (6} \times \text{6 matrix)}$$

$$T_{31} = 0 \quad (6 \times 6 \text{ matrix})$$

$$T_{32} = \begin{bmatrix} \lambda_b [I] & | & 0 \\ \hline 0 & | & \lambda'_b [I] \end{bmatrix}$$

$$T_{11} = A - T_{21}$$

$$T_{22} = A - T_{12} - T_{32}$$

$$\text{and } T_{33} = A - T_{23}$$

where

$$A = \begin{bmatrix} -(2\lambda_a + n) & \mu_a & 0 & m & 0 & 0 \\ 2\lambda_a & -(\mu_a + \lambda_a + n) & 2\mu_a & 0 & m & 0 \\ 0 & \lambda_a & -(2\mu_a + n) & 0 & 0 & m \\ n & 0 & 0 & -2(\lambda'_a + m) & \mu_a & 0 \\ 0 & n & 0 & 2\lambda'_a & -(\mu_a + \lambda'_a + m) & 2\mu_a \\ 0 & 0 & n & 0 & \lambda'_a & -(2\mu_a + m) \end{bmatrix}$$

$B = A$ column matrix with zero entries

and

$$P = \{P_{(0,0)N}, P_{(0,1)N}, \dots, P_{(2,0)S}, P_{(2,1)S}, P_{(2,2)S}\}$$

In the triple subscript of p the first two terms inside the parenthesis indicate the state of the components in subsystems B and A respectively, e.g. (0,1) denotes no component down in subsystem B and one component down in subsystem A . The subscript outside the parenthesis indicates the environment state. Any seventeen equations of (5.9) with

$$\sum_{i=0}^2 \sum_{j=0}^2 \sum_{k=N,S} P_{(i,j)k} = 1.0$$

Table 5.1 Comparison of the results obtained from the reduced model and the original model of the system shown in Fig. 1

Description of the lumped state	Availability of the lumped state	Number of states lumped in each environment state	Availability of one identical state	Availability of one identical state obtained from the non-reduced model
(1)	(2)	(3)	(4)	(5)
(0,0)	0.997155	1	0.997155	0.997156
(0,1)	0.567732×10^{-3}	2	0.283866×10^{-3}	0.283866×10^{-3}
(0,2)	0.172580×10^{-6}	1	0.172580×10^{-6}	0.172580×10^{-6}
(1,0)	0.227221×10^{-2}	2	0.113610×10^{-2}	0.113611×10^{-2}
(1,1)	0.233858×10^{-5}	4	0.584645×10^{-6}	0.584645×10^{-6}
(1,2)	0.203460×10^{-8}	2	0.101730×10^{-8}	0.101729×10^{-8}
(2,0)	0.212768×10^{-5}	1	0.212768×10^{-5}	0.212769×10^{-5}
(2,1)	0.616139×10^{-8}	2	0.308069×10^{-8}	0.308070×10^{-8}
(2,2)	0.961824×10^{-11}	1	0.961824×10^{-11}	0.961873×10^{-11}

may be solved to obtain the vector P of steady state availabilities. The results obtained by solving this set of equations are shown in Table 5.1. In column one, the first number inside the parentheses indicates the number of failed components of subsystem A and the second in B . In column (2), the availabilities of the normal and stormy states are merged together. Column (3) gives the number of identical states in each environment state in the original model. The values of column (4) are obtained by dividing the values in column (2) by those in column (3). The results in column (5) are obtained by the analysis of the complete model of 32 states. It can be seen that the results in columns (4) and (5) have only very slight differences due to rounding off. Complete information about the 32 states of the original system can thus be obtained by analyzing the reduced model of 18 states.

State Space Truncation

It has been seen that the state space can be reduced by merging certain groups of states. Another technique is by truncating the state space, i.e. by neglecting the states whose contribution to the measures of system reliability is insignificant. In systems consisting of independent components, the probability of each state can be calculated individually by the product of the individual component probabilities. The states required for determining the reliability measures are selected, their probabilities calculated and the reliability measures

obtained. The system states which make a negligible contribution to the final results can be neglected.

When dependent transition modes are involved, the system state probabilities cannot be obtained directly and the set of differential or linear algebraic equations must be solved depending on whether time specific or steady state solutions are required. Consider first the steady state condition.

The philosophy behind truncation may be understood by examining the following equation for calculating the probability of the i th state

$$p_i = \sum_{k \in X^-} p_k \lambda_{ki} / \sum_{k \in X^-} \lambda_{ik}$$

The contribution to p_i by a state $k \neq i$ is

$$p_k \lambda_{ki} / \sum_{k \in X^-} \lambda_{ik}$$

i.e. the frequency of encountering i from k divided by the total transition rate out of i . Therefore if the states having low probability are deleted, the probability of state i will not be significantly effected. The states have of course to be deleted prior to solving the set of linear equations. The procedure amounts to assuming that the deleted states have a probability equal to zero. Denoting the set of deleted states by X_T , the probability of this subset if there were no truncation is $p_T = \sum_{i \in X_T} p_i$. Since the probability of the rest of the state space is

now one, i.e.

$$\sum_{i \in (X - X_T)} p_i = 1$$

the probability p_T will be distributed over the states $i \in (X - X_T)$ where X is the system state space. If p_T is small, then the probability distribution of the rest of the states will not be significantly affected. The success of the truncation method depends upon selecting low probability states for truncation. The following consideration should be kept in mind while employing truncation.

1. The probability p_i , $i \in X_T$ is less than p_j , $j \in (X - X_T)$. In words, the biggest probability in the truncated subset should be less than the smallest probability of the remaining state space. In systems consisting of two state components, this is not hard to achieve. The state space may be divided into subsets, each subset having states of a certain level of coincident failures. For a system of n identical components there will be $(n+1)$ subsets. These subsets will have the following states:

Subset number	States description
1	All components up
2	One component down
.	.
.	.
$n+1$	n components down

An arbitrary level of truncation should be first selected; for example the states having three or more than three coincident failures can be truncated. The computation can then be repeated by including the next subset, i.e. the states having three coincident failures. If the new values are not significantly different from the previous ones, the computation can be stopped, otherwise one more subset should be included and the computation repeated.

If the units have derated levels or if there are more than one down state, for example, repair and switching states, the same procedure should be adopted with the addition that the states like switching associated with the repair state should also be retained.

In the state space truncation technique, the probabilities of the states adjacent to the truncation boundary are affected the most and the effect decreases when moving away from the boundary.

2. After the states have been truncated, the state transition diagram should be examined to see if the process of truncation has generated any absorbing states. Since the computer program generates only the transition rate matrix, the absorbing states can be located by examining this matrix. An absorbing state will have transitions into it but not out of it. The i th element of the transition rate matrix gives the transition rate from state i to state j . Therefore if the i th row is empty, this means that the i th state is absorbing. Either the absorbing state should be deleted or the states where truncation has generated this absorbing state should be retained.

Example: Consider a transmission system consisting of five links subjected to a common two state fluctuating environment, normal and stormy weather are 32 states in each environment and a total of 64 states. The distribution of states in each environment state is shown below.

Subset number	1	2	3	4	5	6
Number of elements						
failed	0	1	2	3	4	5
Number of states	1	5	10	10	5	1

If, however, the probabilities of failure of three or more components can be assumed to be zero, the states in subsets 4–6 can be ignored and the matrix of transition rates is reduced from 64×64 to 32×32 . A number of studies were

Sequential truncation can be described as the process of building the reliability model by adding components or subsystems one by one and deleting the low probability states at each step. This method consumes more computation time than direct state space truncation but it is more manageable. In direct truncation, the decision to delete states has to be made prior to the solution of the state probabilities. In sequential truncation, the state probabilities are calculated at each step and the states with probabilities less than a reference value are deleted. The assumption, which is generally valid is that the probability of a given state will be decreased after another component has been added to the system. Assume that at a particular step the system has q states designated $S_1, S_2, \dots, S_{q_i}$ and a component having p states, designated $C_1, C_2, \dots, C_p$ is added. If this component were independent from the system which has been built up to this point, then there would be $q \times p$ states in the resulting system, there being p states for every state of the system up to this point. This may be represented as

$$\begin{array}{cccc} S_1C_1 & S_1C_2 & \dots & S_1C_p \\ S_2C_1 & S_2C_2 & \dots & S_2C_p \\ S_3C_1 & S_3C_2 & \dots & S_3C_p \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ S_qC_1 & S_qC_2 & \dots & S_qC_p \end{array}$$

Table 5.2 Studies of the sensitivity of various systems to be conditional truncation of system states

Number of identical components		= 5	Availability of one identical state in the group					
Average failure rate		= 0.5/year	Failures during S.W. = 20					
Normal weather mean duration		= 200 hours	Failures during S.W. = 80					
Stormy weather mean duration		= 1.5 hours	Failures during S.W. = 200					
A. Availabilities of all possible states								
Group number down	Number of identical components in the group	Number of identical states in the group	$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours	$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours
1	0	1	0.998576	0.997154	0.993611	0.998598	0.997203	0.993729
2	1	5	0.284518 $\times 10^{-3}$	0.568107 $\times 10^{-3}$	0.127351 $\times 10^{-2}$	0.277285 $\times 10^{-3}$	0.551913 $\times 10^{-3}$	0.123463 $\times 10^{-2}$
3	2	10	0.173558 $\times 10^{-6}$	0.531820 $\times 10^{-6}$	0.213447 $\times 10^{-5}$	0.158411 $\times 10^{-5}$	0.366977 $\times 10^{-5}$	0.960027 $\times 10^{-5}$
4	3	10	0.354657 $\times 10^{-9}$	0.123808 $\times 10^{-8}$	0.595186 $\times 10^{-8}$	0.181082 $\times 10^{-7}$	0.517200 $\times 10^{-7}$	0.157760 $\times 10^{-6}$
5	4	5	0.122932 $\times 10^{-11}$	0.535826 $\times 10^{-11}$	0.287183 $\times 10^{-10}$	0.262773 $\times 10^{-9}$	0.985960 $\times 10^{-9}$	0.368171 $\times 10^{-8}$
6	5	1	0.509552 $\times 10^{-14}$	0.302236 $\times 10^{-13}$	0.198045 $\times 10^{-12}$	0.439865 $\times 10^{-11}$	0.227601 $\times 10^{-10}$	0.108524 $\times 10^{-9}$

B: Percentage difference from the exact values, $MC = 4$

Group number	Number of components down	Number of identical states in the group	% Difference from the exact values in Table A					
			% Failures during S.W. = 20			% Failures during S.W. = 80		
			$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours	$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours
1	0	1	0.0	0.0	0.0	0.0	0.0	0.0
2	1	5	0.0	0.0	0.0	0.0	0.0	0.0
3	2	10	0.0	0.0	0.0	0.000001	0.000001	0.0
4	3	10	0.000013	0.000005	0.000009	0.000036	0.000050	0.000009
5	4	5	0.004525	0.000022	0.001343	0.002240	0.005604	0.001617

C: Percentage difference from the exact values, $MC = 3$

Group number	Number of components down	Number of identical states in the group	% Difference from the exact values in Table A					
			% Failures during S.W. = 20			% Failures during S.W. = 80		
			$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours	$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours
1	0	1	0.0	0.0	0.0	0.0	0.0	0.000002
2	1	5	0.0	0.0	0.0	0.000004	0.000001	0.000003
3	2	10	0.000009	0.000003	0.000003	0.000165	0.000035	0.000042
4	3	10	0.002385	0.001277	0.000535	0.007575	0.004306	0.005267

D: Percentage difference from the exact values, $MC = 2$

Group number	Number of components down	Number of identical states in the group	% Difference from the exact values in Table A					
			% Failures during S.W. = 20			% Failures during S.W. = 80		
			$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours	$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours
1	0	1	0.0	0.000001	0.000006	0.000018	0.000052	0.000159
2	1	5	0.000001	0.000002	0.000001	0.000065	0.000062	0.000175
3	2	10	0.000466	0.000557	0.000928	0.007762	0.009008	0.008023

E: Percentage difference from the exact values, $MC = 1$

Group number	Number of components down	Number of identical states in the group	% Difference from the exact values in Table A					
			% Failures during S.W. = 20			% Failures during S.W. = 80		
			$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours	$\gamma = 5$ hours	$\gamma = 10$ hours	$\gamma = 22.5$ hours
1	0	1	0.000174	0.000534	0.002139	0.001583	0.003695	0.009718
2	1	5	0.000226	0.000194	0.002276	0.014940	0.013217	0.016551

The probability of state $S_i C_j$ will be the product of the probabilities of states S_i and C_j . Since the probability of state C_j cannot be greater than one, the probability of $S_i C_j$ will be always less than the probability of S_i . Therefore if the probability of state S_i is less than the reference value, the probability of $S_i C_j$ will also be less than the reference value.

When, however, dependent transition modes are involved, some of the above stage combinations may not exist. The probabilities of the resulting states will, however, be generally less than those of the states prior to combination. The method of sequential truncation will be illustrated by application to an example system.

The system consists of three identical railway stations as shown in Fig. 5.9. Each station has a station lane on which platform facilities are provided for passenger disembarkation and a through expressway without any platform facilities. When the station lane and through expressway are both down, traffic cannot pass through and the system is considered failed. It is assumed that under this condition, no further component failure takes place. The guideway is assumed to be perfectly reliable. The results of this subsystem are to be combined with the other subsystems and therefore the probabilities and frequencies of various states of this system are to be determined. The failure and repair rate of the station lane are denoted by λ_s and μ_s and those of the expressway by λ_e and μ_e . The following numerical values have been used

$$\text{Mean Up Time of the station lane} = \frac{1}{\lambda_s} = 800 \text{ Hours}$$

$$\text{Mean Down Time of the station lane} = \frac{1}{\mu_s} = 2 \text{ Hours}$$

$$\text{Mean up time of the expressway} = \frac{1}{\lambda_e} = 1000 \text{ Hours}$$

$$\text{Mean down time of the expressway} = \frac{1}{\mu_e} = 2 \text{ hours}$$

As the guideway is considered perfectly reliable, it can be left out of the analysis. The state transition diagram of a station is shown in Fig. 5.10, U stands for the up state, i.e. when both lanes are up. D means that the station

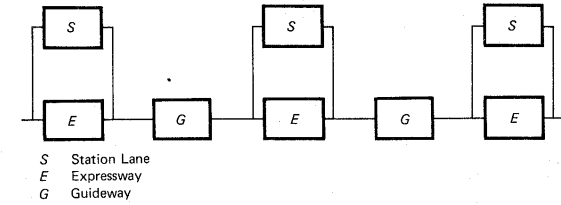


Fig. 5.9. The subsystem functional diagram

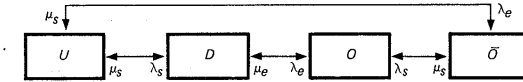


Fig. 5.10 The state transition diagram of a station

lane is down but the traffic can pass through the expressway. O denotes the complete station outage, i.e. both lanes are out and O stands for partial outage, i.e. the station lane is working but the through expressway is down. These states can be represented in the computer as shown in Table 5.3A. The addition of one more station is shown in Table 5.3B. For each state of the component, there is the set of system states of Table 5.3A except that state (3,3) is an impossible state since it means that the two stations are completely out. This is not possible as the exposure to failure is reduced to zero as soon as one station is completely out. The system states in Table 5.3B are numbered in the serial order. The numbers in brackets indicate the combination, the first number indicating the state number of the system before addition and the second indicating the state number of the component being added. Identical states can now be grouped together and the resulting description is shown in Table 5.3C. The states with probabilities less than 10^{-5} are now deleted and the resulting description is given in Table 5.3D. The state numbers are the serial numbers and have no relationship to the state numbers in Table 5.3C.

When the third station is added, the resulting states are shown in Table 5.3E and the states after merging identical states are given in Table 5.3F. The state probabilities are also indicated. If more stations are to be added, then the states with probabilities less than 10^{-5} can again be deleted and the procedure repeated. The exact results, i.e. without any truncation, are shown in Table 5.4

and it can be seen that the results are almost identical. In general, the results are slightly affected depending on the reference probability value employed for truncation.

Table 5.3

A. Model of a single station

System state	Identical states from B	Number of stations in state				Probability
		U	D	0	$\bar{0}$	
1		1	0	0	0	
2		0	1	0	0	
3		0	0	1	0	
4		0	0	0	1	

B. Addition of a station

1	(1,1)	2	0	0	0	
2	(2,1)	1	1	0	0	
3	(3,1)	1	0	1	0	
4	(4,1)	1	0	0	1	
5	(1,2)	1	1	0	0	
6	(2,2)	0	2	0	0	
7	(3,2)	0	1	1	0	
8	(4,2)	0	1	0	1	
9	(1,3)	1	0	1	0	
10	(2,3)	0	1	1	0	
11	(4,3)	0	0	1	1	
12	(1,4)	1	0	0	1	
13	(2,4)	0	1	0	1	
14	(3,4)	0	0	1	1	
15	(4,4)	0	0	0	2	

C. Merging of identical states

1	1	2	0	0	0	0.991051 $\times 10^0$
2	2,5	1	1	0	0	0.495525 $\times 10^{-2}$
3	3,9	1	0	1	0	0.992536 $\times 10^{-5}$
4	4,12	1	0	0	1	0.396420 $\times 10^{-2}$
5	6	0	2	0	0	0.618994 $\times 10^{-5}$
6	7,10	0	1	1	0	0.165058 $\times 10^{-7}$
7	8,13	0	1	0	1	0.990308 $\times 10^{-5}$
8	11,14	0	0	1	1	0.132036 $\times 10^{-7}$
9	15	0	0	0	2	0.396090 $\times 10^{-5}$

D. Truncation of states with probabilities less than 10^{-5}

System state	Number of stations in state				Probability
	U	D	0	$\bar{0}$	
1	2	0	0	0	0.991051 $\times 10^0$
2	1	1	0	0	0.495525 $\times 10^{-2}$
3	1	0	1	0	0.992536 $\times 10^{-5}$
4	1	0	0	1	0.396420 $\times 10^{-2}$
5	0	2	0	0	0.618994 $\times 10^{-5}$
6	0	1	0	1	0.990308 $\times 10^{-5}$
7	0	0	0	2	0.396090 $\times 10^{-5}$

E. Addition of the third station

System state	Number of stations in state				
	U	D	0	$\bar{0}$	
1 (1,1)	3	0	0	0	
2 (2,1)	2	1	0	0	
3 (3,1)	2	0	1	0	
4 (4,1)	2	0	0	1	
5 (5,1)	1	2	0	0	
6 (6,1)	1	1	0	1	
7 (7,1)	1	0	0	2	
8 (1,2)	2	1	0	0	
9 (2,2)	1	2	0	0	
10 (3,2)	1	1	1	0	
11 (4,2)	1	1	0	1	
12 (5,2)	0	3	0	0	
13 (6,2)	0	2	0	1	
14 (7,2)	0	1	0	2	
15 (1,3)	2	0	1	0	
16 (2,3)	1	1	1	0	
17 (4,3)	1	0	1	1	(3,3) not possible
18 (5,3)	0	2	1	0	
19 (6,3)	0	1	1	1	
20 (7,3)	0	0	1	2	
21 (1,4)	2	0	0	1	
22 (2,4)	1	1	0	1	
23 (3,4)	1	0	1	1	
24 (4,4)	1	0	0	2	
25 (5,4)	0	2	0	1	
26 (6,4)	0	1	0	2	
27 (7,4)	0	0	0	3	

F. Merging of identical states

System state	Identical states from E	Number of stations in state				Probability
		U	D	0	$\bar{0}$	
1	1	3	0	0	0	0.986606
2	2,8	2	1	0	0	0.739954×10^{-2}
3	3,15	2	0	1	0	0.148435×10^{-4}
4	4,21	2	0	0	1	0.591964×10^{-2}
5	5,9	1	2	0	0	0.184865×10^{-4}
6	6,11,22	1	1	0	1	0.295760×10^{-4}
7	7,24	1	0	0	2	0.118294×10^{-4}
8	10,16	1	1	1	0	0.493508×10^{-7}
9	12	0	3	0	0	0.153901×10^{-7}
10	13,25	0	2	0	1	0.369310×10^{-7}
11	14,26	0	1	0	2	0.295407×10^{-7}
12	17,23	1	0	1	1	0.394773×10^{-7}
13	18	0	2	1	0	0.461670×10^{-10}
14	19	0	1	1	1	0.738569×10^{-10}
15	20	0	0	1	2	0.295387×10^{-10}
16	27	0	0	0	3	0.787643×10^{-8}

Table 5.4 *Model of three stations without truncation*

System state	State number as in 5.3F	Number of stations in state				Probability
		U	D	0	$\bar{0}$	
1	1	3	0	0	0	0.986606
2	2	2	1	0	0	0.739954×10^{-2}
3	3	2	0	1	0	0.148435×10^{-4}
4	4	2	0	0	1	0.591964×10^{-2}
5	5	1	2	0	0	0.184865×10^{-4}
6	8	1	1	1	0	0.493508×10^{-7}
7	6	1	1	0	1	0.295760×10^{-4}
8	12	1	0	1	1	0.394773×10^{-7}
9	7	1	0	0	2	0.118294×10^{-4}
10	9	0	3	0	0	0.153901×10^{-7}
11	13	0	2	1	0	0.461824×10^{-10}
12	10	0	2	0	1	0.369310×10^{-7}
13	14	0	1	1	1	0.738538×10^{-10}
14	11	0	1	0	2	0.295407×10^{-7}
15	Deleted	0	1	2	0	0.123121×10^{-12}
16	Deleted	0	0	2	1	0.984465×10^{-13}
17	15	0	0	1	2	0.295264×10^{-10}
18	16	0	0	0	3	0.787643×10^{-8}

References

1. R. Billinton and C. Singh, Reliability Evaluation in Large Transmission Systems, Paper No. C 72 475-2, *IEEE Summer Power Meeting* (1972).
2. C. Singh, Reliability Modelling and Evaluation in Electric Power Systems, Ph.D. Thesis, University of Saskatchewan, Saskatoon (August 1972).
3. C. Singh and R. Billinton, A New Method to Determine the Failure Frequency of a Complex System, *Micro-Electronics and Reliability*, Vol. 12, No. 5 (1974).

CHAPTER 6

Reliability Modelling in Non-Markovian Systems

Introduction

Most reliability models assume that the up and down times of the components are exponentially distributed. This assumption leads to a Markovian model with constant interstate transition rates. The analysis in such cases is relatively simple and the numerical results can be easily obtained. The assumption is often valid for the up time but the down times are likely to have a non-exponential distribution. When the components are independent, the steady state results, as is shown later, are not affected by the shape of the distribution. In the case of dependent failures there can be a very definite affect.

If the distributions cannot be represented by a single exponential form then the process becomes non-Markovian and different techniques are required for system solution. This chapter presents some different methods for solving non-Markovian systems by application to specific models.

The Difficulty with Non-Markovian Processes

The essential difficulty with non-Markovian processes can be illustrated by the analysis of the reliability model of a binary unit (See Fig. 2.8). The up and down time durations are assumed to have the distribution $\lambda \exp\{-\lambda x\}$ and $f(y)$ respectively. Denoting the state occupied at time t by $Z(t)$, the equation of state 0, i.e. the up state can be written as

$$P\{Z(t + \Delta t) = 0\} = P\{Z(t + \Delta t) = 0 | Z(t) = 0\} P\{Z(t) = 0\} + P\{Z(t + \Delta t) = 0 | Z(t) = 1\} P\{Z(t) = 1\} \quad (6.1)$$

It can be seen that as $\Delta t \rightarrow 0+$,

$$P\{Z(t + \Delta t) = 0 | Z(t) = 0\} = 1 - \lambda \Delta t$$

The distribution of down time is, however, not exponential and therefore the repair rate $\mu(y)$ depends on the time y which the component has spent in state 1. The coefficient of the second term on the right hand side of Equation (6.1) under the condition that the component has been in state 1 for time y , is therefore $\mu(y)\Delta t$. The required coefficient can be obtained by integrating this conditional

coefficient over the distribution of the time spent in state 1 up to time t . It is this dependence of the transfer rate on the random duration y in the state 1 that is the essential difficulty in formulating Equation (6.1). If, however, the transition rates were to depend on time t in an explicitly known manner, there would not be any special difficulty. For example, if the repair rate were known as a function of time t , the coefficient of the second term would be $\mu(t)\Delta t$. The basic approach to avoiding this difficulty is to convert the non-Markov process into a Markov process by redefinition of the state space.

Method of Supplementary Variables

This is probably the most direct method of dealing with non-Markovian systems and will be illustrated by application to a bank of three single phase transformers with one spare. The state transition diagram with unrestricted repair and exponentially distributed down times has been presented in Fig. 4.4. The state

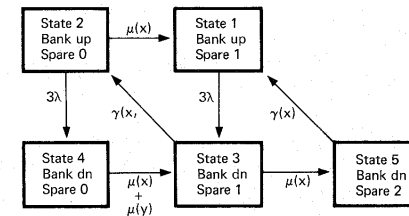


Fig. 6.1. The state transition diagram of a bank of three single phase transformers with one spare, unrestricted repair.

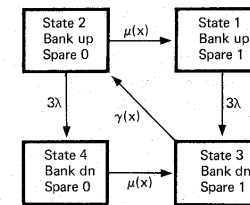


Fig. 6.2. The state transition diagram of a bank of three single phase transformers with one spare, restricted repair.

space diagram with arbitrary down time distributions and unrestricted repair is shown in Fig. 6.1. The corresponding model when the repair facilities are restricted is shown in Fig. 6.2. In this model, the repair on the second unit is not started until the repair on the failed unit is completed and it is reinstalled. The difference between the two state transition diagrams lies in the elimination of state 5 and the transition rate out of state 4. When the repair is unrestricted, both the transformers are simultaneously under repair, one for time x and the other for time y . With the repair restricted, only one transformer is under repair in state 4 and the other failed transformer is waiting. The repair on the second transformer cannot start until the first one has been installed and therefore state 5 is eliminated. The following notation has been used.

$f_u(x), f_r(x), f_c(x)$ = The probability density functions for the up time, repair duration and the change out period respectively

$S_u(x), S_r(x), S_c(x)$ = The survivor functions for the up time, repair duration and the change out period respectively

For example, $S_u(x) = P(U > x) = \int_x^\infty f_u(y) dy$, U is a random variable denoting the up time of the transformer.

Let Y be a continuous positive random variable specifying the duration of a particular state of the component until the termination of that state. Then the age specific transition rate is defined as

$$\phi(y) = \lim_{\Delta y \rightarrow 0^+} \frac{P(y < Y \leq y + \Delta y | y < Y)}{\Delta y} = \frac{f(y)}{S(y)}$$

$\mu(x)$ = The repair rate when the repair has been going on for x period of time

$\gamma(x)$ = The change out rate when the change out has been in operation for time x .

and λ = The failure rate of the transformer. This is constant as the distribution for the up time is assumed to be exponential.

Exponential Models

If the repair and change out times are also assumed to be exponentially distributed, then the transition rates are all constant, i.e. $\mu(x) = \mu$ and $\gamma(x) = \gamma$. The process is Markovian as the random variables that generate it are all exponentially distributed.

Unrestricted Repair

This case has already been discussed in Chapter 4.

Restricted Repair

The equations for the state transition diagram of Fig. 6.2 can be written as

$$3\lambda p_1 - \mu p_2 = 0$$

$$(3\lambda + \mu)p_2 - \gamma p_3 = 0$$

$$\mu p_4 - 3\lambda p_2 = 0$$

$$\gamma p_3 - 3\lambda p_1 - \mu p_4 = 0$$

Any three equations of the above set together with the relationship

$$\sum_{i=1}^4 p_i = 1$$

can be solved to obtain the steady state availabilities

$$p_1 = 1/Z \quad p_2 = \left(\frac{3\lambda}{\mu}\right)/Z$$

$$p_3 = \left(\frac{3\lambda + \mu}{\gamma}\right)\frac{3\lambda}{\mu}/Z \quad p_4 = \left(\frac{3\lambda}{\mu}\right)^2/Z$$

where

$$Z = 1 + \frac{3\lambda}{\mu} \left(1 + \frac{3\lambda + \mu}{\gamma} + \frac{3\lambda}{\mu}\right)$$

$$P_{DN} = \frac{3\lambda}{\mu} \left(\frac{3\lambda}{\mu} + \frac{3\lambda + \mu}{\gamma}\right)/Z$$

and

$$f_{DN} = f_{UP} = 3\lambda \left(1 + \frac{3\lambda}{\mu}\right)/Z$$

Non-exponential Models

A. General Distributions for Repair and Change Out Periods, Restricted Repair.

When the repair and change out period distributions are non-exponential, the repair rate, $\mu(x)$ and the change out rate $\gamma(z)$ are functions of the age of the repair and change out. The stochastic process is, therefore, non-Markovian. The

most direct method of tackling this process is by the inclusion of sufficient supplementary variables in the specification of the state of the system to make the process Markovian. In the case of transformer banks the supplementary variables are the times expended in the repair or change out process. The resulting Markov process is in continuous time and has a state space which is multi-dimensional and partly discrete and partly continuous.

Define

$p_i(t)$ = The probability that the system is in state i at time t .

$$p_i(x; t) = \lim_{\Delta x \rightarrow 0+} \frac{P[\text{The system is in state } i, \text{ the current operation having been started in } (t-x-\Delta x, t-x)]}{\Delta x}$$

The current operation is the process of repair or change out due to which the system is in state i and as soon as this operation is terminated the system will transit out of this state. For example in Fig. 6.2

$p_4(x; t) \Delta x$ = The probability that the system is in state 4 at time t and the elapsed time since the repair started on the transformer bank lies in the interval $(x, x+\Delta x)$

The forward equations of the resulting Markov process can be written by considering the transitions during the increment Δt .

$$p_1(t + \Delta t) = p_1(t)(1 - 3\lambda\Delta t) + \Delta t \int_0^\infty p_2(x; t)\mu(x)dx$$

$$p_2(x + \Delta t; t + \Delta t) = p_2(x; t) \{1 - (\mu(x) + 3\lambda)\Delta t\}$$

$$p_3(x + \Delta t; t + \Delta t) = \{1 - \gamma(x)\Delta t\} p_3(x; t)$$

$$p_4(x + \Delta t; t + \Delta t) = \{1 - \mu(x)\Delta t\} p_4(x; t) + 3\lambda\Delta t p_2(x; t)$$

The resulting differential equations as $\Delta t \rightarrow 0+$ are

$$\frac{\partial p_1(t)}{\partial t} = -3\lambda p_1(t) + \int_0^\infty p_2(x; t)\mu(x)dx \quad (6.2)$$

$$\frac{\partial p_2(x; t)}{\partial t} + \frac{\partial p_2(x; t)}{\partial x} = -\{3\lambda + \mu(x)\} p_2(x; t) \quad (6.3)$$

$$\frac{\partial p_3(x; t)}{\partial t} + \frac{\partial p_3(x; t)}{\partial x} = -\gamma(x) p_3(x; t) \quad (6.4)$$

$$\frac{\partial p_4(x; t)}{\partial t} + \frac{\partial p_4(x; t)}{\partial x} = -\mu(x) p_4(x; t) + 3\lambda p_2(x; t) \quad (6.5)$$

Equations (6.2)–(6.5) can be solved under the boundary conditions

$$p_2(0; t) = \int_0^\infty p_3(x; t)\gamma(x)dx \quad (6.6)$$

$$p_3(0; t) = 3\lambda p_1(t) + \int_0^\infty p_4(x; t)\mu(x)dx \quad (6.7)$$

and

$$p_4(0; t) = 0 \quad (6.8)$$

The boundary Equation (6.6) results from the fact that as soon as the spare or repaired transformer is reinstalled the system enters state 2. Similar reasoning holds for Equation (6.7). Equation (6.8) states that it is impossible to be in state 4 without the transformer being under repair for some time.

It is interesting to indicate at this point that Equations (6.2) – (6.8) can also be obtained using the frequency balancing concept outlined in Chapter 3. Following the approach given in this chapter, the equation for state 2 under the condition that repair has been in progress for time x can be written as

$$\begin{aligned} \Delta \{p_2(x; t) \Delta x\} &= -\{3\lambda + \mu(x)\} p_2(x; t) \Delta x \cdot \Delta t \\ \text{i.e.} \quad \Delta p_2(x; t) &= -\{3\lambda + \mu(x)\} p_2(x; t) \Delta t \end{aligned}$$

Now knowing that small increases in both x and t are Δt

$$\begin{aligned} \Delta p_2(x; t) &= \frac{\partial p_2(x; t)}{\partial t} \Delta t + \frac{\partial p_2(x; t)}{\partial x} \Delta t \\ &= -\{3\lambda + \mu(x)\} p_2(x; t) \Delta t \end{aligned}$$

That is

$$\frac{\partial p_2(x; t)}{\partial t} + \frac{\partial p_2(x; t)}{\partial x} = -\{3\lambda + \mu(x)\} p_2(x; t)$$

which is the same as Equation (6.3).

Since the primary interest is in the steady state availabilities, Equations (6.2)–(6.8) reduce to the equilibrium equations as $t \rightarrow \infty$

$$3\lambda p_1 = \int_0^\infty p_2(x)\mu(x)dx \quad (6.9)$$

$$\frac{\partial p_2(x)}{\partial x} = -\{3\lambda + \mu(x)\} p_2(x) \quad (6.10)$$

$$\frac{\partial p_3(x)}{\partial x} = -\gamma(x) p_3(x) \quad (6.11)$$

$$\frac{\partial p_4(x)}{\partial x} = -\mu(x)p_4(x) + 3\lambda p_2(x) \quad (6.12)$$

$$p_2(0) = \int_0^\infty p_3(x)\gamma(x)dx \quad (6.13)$$

$$p_3(0) = 3\lambda p_1 + \int_0^\infty p_4(x)\mu(x)dx \quad (6.14)$$

$$\text{and } p_4(0) = 0 \quad (6.15)$$

In these equations

p_i = The steady state probability of being in state i

and

$p_i(x)\Delta x$ = The steady state probability of being in state i and the elapsed time since the current operation started lies in the interval $(x, x + \Delta x)$.

It should be noted that $p_i(x)\Delta x$ denotes the probability of a continuous state since x may lie anywhere in the interval $(0, \infty)$. The availability of the system condition denoted by state i can, therefore, be obtained by

$$p_i = \int_0^\infty p_i(x)dx \quad (6.16)$$

Equations (6.9)–(6.15) together with the normalizing equation

$$\sum_{i=1}^4 p_i = 1 \quad (6.17)$$

can be solved to obtain the steady state availabilities.

Equation (6.11) on solution gives

$$p_3(x) = p_3(0) \exp \left\{ -\int_0^x \gamma(w)dw \right\} \quad (6.18)$$

Since it is well known that

$$S_c(x) = \exp \left\{ -\int_0^x \gamma(w)dw \right\}$$

$$p_3 = \int_0^\infty p_3(x)dx = p_3(0) \int_0^\infty S_c(x)dx = p_3(0)/\gamma$$

where

$$\frac{1}{\gamma} = \text{the mean change out period}$$

Therefore

$$p_3 = \frac{1}{\gamma} \left\{ 3\lambda p_1 + \int_0^\infty p_4(x)\mu(x)dx \right\} \quad (6.19)$$

Solving Equation (6.10)

$$p_2(x) = p_2(0) \exp \left\{ -\int_0^x (3\lambda + \mu(w))dw \right\} \quad (6.20)$$

$$p_2 = \int_0^\infty p_2(x)dx = p_2(0) \int_0^\infty \exp(-3\lambda x) S_r(x)dx$$

$$= \frac{1}{3\lambda} (1-E) \int_0^\infty p_3(x)\gamma(x)dx \quad (6.21)$$

where

$$E = \int_0^\infty f_r(x) e^{-3\lambda x} dx = E_r(e^{-3\lambda x}) \quad (6.22)$$

Substituting from Equation (6.18) into (6.20)

$$p_2 = \frac{1}{3\lambda} (1-E) p_3 \gamma \quad (6.23)$$

Substituting (6.20) into (6.9)

$$3\lambda p_1 = p_2(0) \int_0^\infty \exp \left\{ -\int_0^x (3\lambda + \mu(w))dw \right\} \mu(x)dx$$

$$= p_2(0) \int_0^\infty f_r(x) e^{-3\lambda x} dx = \gamma E p_3 \quad (6.24)$$

Equation (6.12) can be solved using the boundary condition (6.15) giving

$$p_4(x) = \exp \left\{ -\int_0^x \mu(w)dw \right\} \int_0^x 3\lambda p_2(y) \exp \left\{ \int_0^y \mu(v)dv \right\} dy$$

Substituting from Equation (6.20) and simplifying

$$p_4(x) = \gamma p_3 (1 - e^{-3\lambda x}) S_r(x) \quad (6.25)$$

and

$$p_4 = \int_0^\infty p_4(x)dx = \gamma p_3 \left[\frac{1}{\mu} - \frac{(1-E)}{3\lambda} \right] \quad (6.26)$$

where

$$\frac{1}{\mu} = \text{Mean repair time}$$

From (6.24)

$$p_3 = \frac{3\lambda}{\gamma E} p_1 \quad (6.27)$$

From (6.23)

$$p_2 = \frac{1-E}{3\lambda} \gamma p_3 = \frac{1-E}{E} p_1 \quad (6.28)$$

From (6.26) and (6.27)

$$p_4 = \frac{3\lambda}{E} \left[\frac{1}{\mu} - \frac{1}{3\lambda} (1-E) \right] p_1 \quad (6.29)$$

Substituting from Equations (6.27)–(6.29) into (6.17) and simplifying

$$\begin{aligned} p_1 &= 1/D \quad \text{where} \quad D = 1 + \frac{3\lambda}{E} \left(\frac{1}{\mu} + \frac{1}{\gamma} \right) \\ p_2 &= \frac{1-E}{DE} \quad p_3 = \frac{3\lambda}{\gamma DE} \quad \text{and} \quad p_4 = \frac{3\lambda}{DE} \left[\frac{1}{\mu} - \frac{1-E}{3\lambda} \right] \\ p_{DN} &= p_3 + p_4 = \frac{3\lambda}{DE} \left[\frac{1}{\gamma} + \frac{1}{\mu} - \frac{1-E}{3\lambda} \right] \end{aligned} \quad (6.30)$$

The frequency

$$\begin{aligned} f_{DN} &= \int_0^\infty p_3(x) \gamma(x) dx = p_3(0) \int_0^\infty f_c(x) dx \\ &= p_3(0) = \gamma p_3 = \frac{3\lambda}{DE} \\ f_{UP} &= 3\lambda(p_1 + p_2) = \frac{3\lambda}{DE} = f_{DN} \end{aligned} \quad (6.31)$$

It can be seen that the equations using a general distribution are different from those derived assuming an exponential distribution. The availabilities are now dependent upon the transform E which for the exponential distribution of repair reduces to

$$= \int_0^\infty \mu e^{-(3\lambda+\mu)x} dx = \frac{\mu}{3\lambda + \mu}$$

It is interesting to note that the steady state probabilities depend only upon the reciprocal of the mean change out time, i.e. the average change out rate. Therefore as long as the mean change out time stays the same, the form of its probability density function does not affect the steady state probabilities and frequencies (see that $f_{DN} = p_3 \gamma$). This always happens when the operation is started and completed in the same state as for example in the present case the reinstallation is started and completed in state 3 of the system. If, however, the reinstallation were not always completed in state 3, then the form of the probability density affects the steady state probabilities. This is seen in the next section.

B. General Distribution for Change Out, Exponential Distribution for Repair, Unrestricted Repair.

It can be seen from Fig. 6.1 that in this case the reinstallation phase initiated in state 3 may not always be completed in state 3 but sometimes in state 5. The concept of average change out rate thus does not apply here.

Defining

$$p_i(x; t) \Delta x = \text{The probability that the system is in the state } i \text{ at time } t \text{ and the elapsed time since the start of reinstallation lies in the interval } (x, x + \Delta x)$$

the differential equations can be written as

$$\frac{\partial p_1(t)}{\partial t} = -3\lambda p_1(t) + \mu p_2(t) + \int_0^\infty p_5(x; t) \gamma(x) dx \quad (6.32)$$

$$\frac{\partial p_2(t)}{\partial t} = -(3\lambda + \mu) p_2(t) + \int_0^\infty p_3(x; t) \gamma(x) dx \quad (6.33)$$

$$\frac{\partial p_3(x; t)}{\partial t} + \frac{\partial p_3(x; t)}{\partial x} = -\{\mu + \gamma(x)\} p_3(x; t) \quad (6.34)$$

$$\frac{\partial p_4(t)}{\partial t} = -2\mu p_4(t) + 3\lambda p_2(t) \quad (6.35)$$

and

$$\frac{\partial p_5(x; t)}{\partial t} + \frac{\partial p_5(x; t)}{\partial x} = -\gamma(x) p_5(x; t) + \mu p_3(x; t) \quad (6.36)$$

These equations can be solved using the boundary conditions

$$p_3(0; t) = 3\lambda p_1(t) + 2\mu p_4(t) \quad (6.37)$$

and

$$p_5(0; t) = 0 \quad (6.38)$$

Equation (6.38) indicates that the reinstallation phase is never started in state 5.

Under equilibrium conditions, i.e. as $t \rightarrow \infty$, the Equations (6.32)–(6.38) reduce to

$$3\lambda p_1 = \mu p_2 + \int_0^\infty p_5(x) \gamma(x) dx \quad (6.39)$$

$$(3\lambda + \mu) p_2 = \int_0^\infty \gamma(x) p_3(x) dx \quad (6.40)$$

$$\frac{\partial p_3(x)}{\partial x} = -\{\mu + \gamma(x)\} p_3(x) \quad (6.41)$$

$$2\mu p_4 = 3\lambda p_2 \quad (6.42)$$

$$\frac{\partial p_5(x)}{\partial x} = -\gamma(x) p_5(x) + \mu p_3(x) \quad (6.43)$$

$$p_3(0) = 3\lambda p_1 + 2\mu p_4 \quad (6.44)$$

and

$$p_5(0) = 0 \quad (6.45)$$

On solving Equation (6.41) and substituting from Equation (6.44)

$$p_3(x) = (3\lambda p_1 + 2\mu p_4) \exp \left\{ -\int_0^x (\mu + \gamma(w)) dw \right\} \quad (6.46)$$

Therefore

$$\begin{aligned} p_3 &= \int_0^\infty p_3(x) dx = (3\lambda p_1 + 2\mu p_4) \int_0^\infty e^{-\mu x} S_c(x) dx \\ &= (3\lambda p_1 + 2\mu p_4) \left(\frac{1 - E_c}{\mu} \right) \end{aligned} \quad (6.47)$$

where

$$E_c = \int_0^\infty f_c(x) e^{-\mu x} dx$$

Substituting (6.46) and (6.40) and simplifying

$$(3\lambda + \mu) p_2 = (3\lambda p_1 + 2\mu p_4) E_c \quad (6.48)$$

Solving Equations (6.43) using (6.45)

$$p_5(x) = \exp \left\{ -\int_0^x \gamma(w) dw \right\} \int_0^x p_3(y) \exp \left\{ \int_0^y \gamma(v) dv \right\} dy$$

Substituting from Equation (6.46) and simplifying

$$p_5(x) = p_3(0)(1 - e^{-\mu x}) S_c(x) \quad (6.49)$$

Therefore

$$p_5 = \int_0^\infty p_5(x) dx = (3\lambda p_1 + 2\mu p_4) \left[\frac{1}{\gamma} - \frac{1}{\mu} + \frac{E_c}{\mu} \right] \quad (6.50)$$

Substituting from (6.42) into (6.48) and simplifying

$$p_2 = \frac{3\lambda E_c}{3\lambda(1 - E_c) + \mu} p_1 \quad (6.51)$$

From (6.42)

$$p_4 = \frac{3\lambda}{2\mu} p_2 = \frac{9\lambda^2 E_c}{2\mu[3\lambda(1 - E_c) + \mu]} p_1 \quad (6.52)$$

Substituting for p_4 in (6.47)

$$p_3 = \frac{3\lambda}{\mu} (1 - E_c) \left[1 + \frac{3\lambda E_c}{3\lambda(1 - E_c) + \mu} \right] p_1 \quad (6.53)$$

Substituting for p_4 in (6.50)

$$p_5 = \frac{3\lambda}{\mu} \left(\frac{\mu}{\gamma} - 1 + E_c \right) \left[1 + \frac{3\lambda E_c}{3\lambda(1 - E_c) + \mu} \right] p_1 \quad (6.54)$$

Substituting into the normalizing equation

$$\sum_{i=1}^5 p_i = 1$$

and simplifying

$$p_1 = \frac{2\gamma\mu\{3\lambda(1 - E_c) + \mu\}}{B}$$

where

$$B = \{3\lambda(1 - E_c) + \mu\}(2\gamma\mu + 6\lambda\mu) + 3\lambda E_c(2\gamma\mu + 6\lambda\mu + 3\lambda\gamma)$$

$$p_2 = \frac{6\lambda\gamma\mu}{B} E_c \quad p_3 = \frac{6\lambda\gamma(3\lambda + \mu)}{B} (1 - E_c)$$

$$p_4 = \frac{9\lambda^2\gamma}{B} E_c \quad p_5 = \frac{6\lambda(3\lambda + \mu)(\mu - \gamma + \gamma E_c)}{B}$$

$$p_{DN} = p_3 + p_4 + p_5 = \frac{9\lambda^2\gamma E_c + 6\lambda\mu(3\lambda + \mu)}{B} \quad (6.55)$$

$$p_{UP} = \frac{6\lambda\gamma\mu + 2\gamma\mu^2}{B} \quad (6.56)$$

$$\text{The Frequency} \quad f_{DN} = \int_0^\infty p_3(x) \gamma(x) dx + \int_0^\infty p_5(x) \gamma(x) dx$$

$$\begin{aligned}
&= p_3(0) \int_0^\infty \gamma(x) \exp \left\{ - \int_0^x (\mu + \gamma(w)) dw \right\} dx \\
&\quad + p_3(0) \int_0^\infty (1 - e^{-\mu x}) f_c(x) dx \\
&= p_3(0) = 3\lambda p_1 + 2\mu p_4 = 3\lambda(p_1 + p_2) = 3\lambda p_{UP} = f_{UP} \quad (6.57)
\end{aligned}$$

Semi-Markov Processes

A semi-Markov process is a stochastic process in which the transitions from state to state are in accordance with a Markov Chain but the time spent in a state before a transition occurs is random.

Consider a stochastic process $Z(t)$ which at time t can be in any of the n distinct states, $Z(t) = i$ denoting that the stochastic process is in state i at time t . Let the time just after the m th transition be denoted by t_m . The stochastic process is Markovian if

$$\begin{aligned}
P[Z(t_m) = j | Z(t_{m-1}) = i, Z(t_{m-2}) = k, \dots, Z(t_1) = l] \\
= P[Z(t_m) = j | Z(t_{m-1}) = i]
\end{aligned}$$

If this probability is independent of the number of the transition, the process is time homogeneous.

Let

$$\alpha_{ij} = P[Z(t_m) = j | Z(t_{m-1}) = i]$$

be the probability of going from state i to j in one step. The matrix of these transition probabilities will be denoted by $A = (\alpha_{ij})$. Given that the state i has just been entered, the probability of going to state j is specified by α_{ij} and $F_{ij}(\cdot)$ is the distribution function of the waiting time X_{ij} in state i given that the next transition will be to state j . The transitions can therefore be thought of as taking place in two stages. When the process has just entered state i , the next state j is selected according to the matrix A but once j has been picked, the waiting time X_{ij} is specified by $F_{ij}(\cdot)$. The Markov Chain, $A = (\alpha_{ij})$ associated with the semi-Markov process is called an embedded Markov Chain.

Discrete time and continuous time Markov chains are special cases of the semi-Markov process. For a discrete time Markov chain

$$F_{ij}(t) = \begin{cases} 0 & \text{for } t < c \\ 1 & \text{for } t \geq c \end{cases}$$

where c is a constant. The discrete time Markov chain is therefore a semi-Markov process in which the waiting time X_{ij} is constant. For the continuous time Markov chain

$$F_{ij}(t) = \begin{cases} 0 & \text{for } t \leq 0 \\ 1 - \exp \{-\lambda_i t\} & \text{for } t > 0, \lambda_i > 0 \end{cases}$$

where λ_i is the reciprocal of the mean time of X_{ij} .

Let

$$Q_{ij}(t) = \alpha_{ij} F_{ij}(t)$$

It should be noted that $F_{ij}(t)$ represents the conditional probability that a transition will take place in time $\leq t$ given that the process has just entered i and will next enter j . $Q_{ij}(t)$ on the other hand is the probability that given that the process has just entered i , it will transit to state j in time less than or equal to t . Let

$$p_{ij}(t) = P[Z(t) = j | Z(0) = i]$$

Then

$$p_{ij}(t) = \sum_{k=0}^n \int_0^t p_{kj}(t-x) dQ_{ik}(x) \quad (6.58)$$

and

$$p_{ij}(t) = 1 - \sum_{k=0}^n \int_0^t (1 - p_{kj}(t-x)) dQ_{ik}(x) \quad (6.59)$$

The above equations involve convolution integrals of the form

$$\int_0^t g(t-x) f(x) dx$$

The Laplace of this integral is of the simple form $\bar{g}(s)\bar{f}(s)$. These integral equations can therefore be reduced to linear equations by taking Laplace transforms. The Laplace transform of Equations (6.58) and (6.59) in the matrix form can be written as

$$s\bar{P}(s) = [I - \bar{G}(s)]^{-1} [I - \bar{F}_d(s)] \quad (6.60)$$

where

$\bar{P}(s)$ is the matrix whose $p_{ij}(s)$ term is the Laplace of the probability $p_{ij}(t)$
 $\bar{G}(s)$ is the matrix whose ij^{th} term is $\alpha_{ij} \bar{f}_{ij}(s)$
 $\bar{F}_d(s)$ is the diagonal matrix whose ii^{th} term is equal to $\sum_k \bar{f}_{ik}(s) \alpha_{ik}$
 I is the identity matrix.

The probabilities can be obtained in the Laplace form and the time specific solutions found by inversion. More often, however, the steady state probabilities are required and these can be obtained directly using the following relationship

$$p_i = \frac{m_i \Pi_i}{\sum_j m_j \Pi_j} \quad (6.61)$$

where

π_i = The steady state probability that the embedded Markov chain is in state i .

m_i = The mean residence time in state i .

The steady state probability vector can be found by solving

$$\pi A = \pi$$

along with

$$\sum_i \pi_i = 1$$

The application of a semi-Markov process is illustrated with an example of two three phase transformers in parallel. When a fault develops on either of the transformers, both the transformers are shut down. After the defective transformer has been isolated, the good one is returned for operation. The state transition diagram is shown in Fig. 6.3.

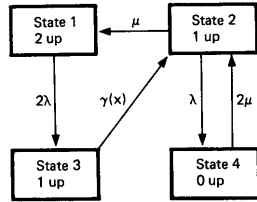


Fig. 6.3 The state transition diagram for two three phase transformers in parallel

The up time and repair time are assumed to be exponentially distributed but the change out time is assumed to have an arbitrary probability density function $f_c(t)$ having a Laplace transform $\bar{f}_c(s)$. The A matrix of transformation probabilities is

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ \frac{\mu}{\mu + \lambda} & 0 & 0 & \frac{\lambda}{\lambda + \mu} \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

The steady state probabilities of the Markov Chain are

$$\pi_1 = \frac{1}{3 + 2\lambda/\mu} \quad \pi_2 = \left(\frac{\mu + \lambda}{\mu}\right) \left/ \left(3 + \frac{2\lambda}{\mu}\right) \right.$$

$$\pi_3 = \frac{1}{3 + 2\lambda/\mu} \quad \pi_4 = \left(\frac{\lambda}{\mu}\right) \left/ \left(3 + \frac{2\lambda}{\mu}\right) \right.$$

Also

$$m_1 = \frac{1}{2\lambda} \quad m_2 = \frac{1}{\mu + \lambda} \quad m_3 = \frac{1}{\gamma} \quad m_4 = \frac{1}{2\mu}$$

where

$$\frac{1}{\gamma} = \text{mean change out time.}$$

Substituting into Equation (6.61)

$$p_1 = \frac{1}{Z} \quad p_2 = \frac{2\lambda}{\mu Z}$$

$$p_3 = \frac{2\lambda}{\gamma Z} \quad p_4 = \frac{\lambda^2}{\mu^2 Z}$$

where

$$Z = 1 + \frac{\lambda^2}{\mu^2} + \frac{2\lambda}{\mu} + \frac{2\lambda}{\gamma}$$

It can be seen that the probability density function of the change out time $f_c(t)$ enters these expressions through the mean value $1/\gamma$. The same results would therefore be obtained if the change out time were assumed to be exponentially distributed with the transition rate equal to the reciprocal of the mean change out time. This is due to the fact that the change out operation starts and ends in the same state, number 3. It should, however, be noted that the time dependent solution would be different from that assuming an exponential distribution.

Device of Stages

The device of stages is a method of representing a non-exponentially distributed state by a combination of stages each of which is exponentially distributed. The method, therefore, represents a non-Markovian model by an equivalent Markovian model which is generally simpler to solve. Any distribution with a rational Laplace transform can, in principle, be represented exactly by a stage combination. Though this may involve complex probabilities associated with the fictitious stages, the probabilities of the actual states of the system are always real. Many distributions not necessarily having a rational Laplace transform can also be reasonably approximated by relatively simple stage combinations. The application of this technique involves the following steps.

1 Selection of a Stage Combination

When the distribution has a rational Laplace transform, the stage combination can be found by examining the roots of this transform. In other probability distribution cases or of directly fitted data cases a suitable guess has to be made. The probability density function and the hazard rate function of a given distribution or data should be examined. A number of simple stage combinations and their characteristics are described later in the chapter. The characteristics of the given distribution or data should be compared with those of the stage combinations, and a suitable combination should be selected. The difference between distributions like gamma, Weibul and lognormal become significant only in the tail regions, their hazard functions are, however, quite distinct. Therefore both the density function and the hazard function should be compared when selecting a proper stage combination.

2 Determination of Parameters

When a stage combination has been selected, the next step is the derivation of its parameters from those of the distribution. This can be done by a moment matching technique which is described in this chapter.

In addition to the constant hazard rate exhibited only by the exponential distribution, there are four basic hazard rate shapes.

1. Increasing hazard rate (Fig. 6.4a)
2. Decreasing hazard rate (Fig. 6.4b)
3. Initial period of decreasing hazard rate followed by increasing rate (Fig. 6.4c)
4. Initial period of increasing hazard rate followed by decreasing rate (Fig. 6.4d).

The combinations discussed in this chapter are capable of generating these shapes. An awareness of this characteristic can be very useful in selecting a proper combination.

General Technique for Deriving the Characteristics of Stage Combinations

The probability density function, the survivor function and the hazard rate function of a given stage combination can be derived in a number of ways but a general technique, often helpful in difficult situations is described below. Let O be the equivalent state of a given stage combination. The transitions from this state are assumed to be terminated in an absorbing state A , as shown in Fig. 6.5. The process is assumed to start in the same stage, into which it would first transit when state O is entered. The time spent in state O is identical with the time from the origin as no transition is made from A to O . Under these conditions

$$\begin{aligned} f_0(x) &= \text{The probability density function of state } O, \text{ i.e., the stage combination} \\ &= \text{The time specific frequency of transiting from } O \text{ to } A \end{aligned}$$

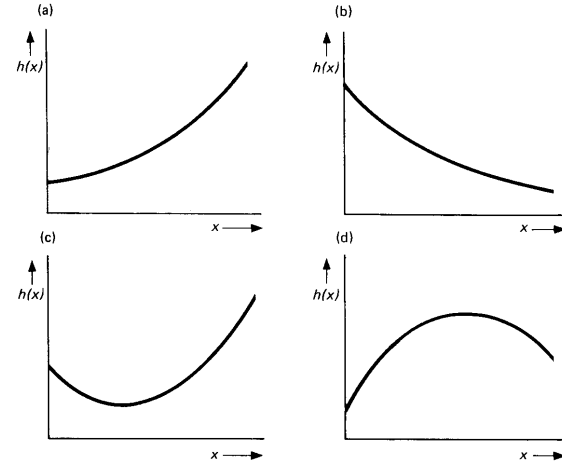


Fig. 6.4 Some typical hazard rate functions.

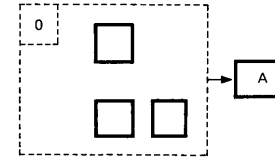


Fig. 6.5 State transition diagram to determine the characteristics of a stage combination

$$= \sum_{i \in 0} p_i(x) \lambda_{iA} \quad (6.62)$$

where

λ_{iA} = The transition rate from state i , member of 0 , to the absorbing state A .

$S_0(x)$ = The survivor function of the stage combination

= The probability of being in state 0 at any time

$$= \sum_{i \in 0} p_i(x) \quad (6.63)$$

and

$$\begin{aligned} \phi_0(x) &= \text{The hazard rate} \\ &= f_0(x)/S_0(x) \\ &= \sum_{i \in 0} p_i(x) \lambda_{iA} / \sum_{i \in 0} p_i(x) \\ &= \text{The equivalent transfer rate from state 0 to A} \end{aligned} \quad (6.64)$$

The evaluation of Equations (6.62) – (6.64) involves the calculation of the probabilities of being in various stages of state O . This is generally a simpler method. The main advantage is that though explicit formulae may not be possible, numerical values of the state probabilities can be easily determined using the techniques discussed in Chapter 3. The numerical values of the three characteristics can then be calculated using the above equations. The method will become more clear later when the specific equations for the stage combinations are derived.

Stage Combinations

We will now describe some specific stage combinations which are simple but versatile enough to approximate a variety of probability distribution functions. The stage combinations will be shown as approximating the down time distribution of a two state component whose up state is assumed to be exponentially distributed. This arrangement is only for the sake of illustration, the stage combination may be used to represent a state in any other context.

The Stages in Series

The stages are traversed in a sequential order and the non-negative continuous random variable X , representing the component state duration is the sum of ' a ' independent exponentially distributed random variables. The Laplace transform of the distribution of X can then easily be proved to be

$$\frac{\rho_1 \dots \rho_a}{(\rho_1 + s) \dots (\rho_a + s)} \quad (6.65)$$

The probability density function can be obtained by expressing Equation (6.64) in the partial fraction form

$$\sum \frac{A_i \rho_i}{\rho_i + s}$$

The probability density function, therefore, is

$$\sum A_i \rho_i \exp(-\rho_i x) \quad (6.66)$$

It should be noted that $\sum A_i = 1$ but A_i do not all lie in $(0, 1)$ and therefore $\{A_i\}$ is not a probability distribution. It can be easily shown that for a fixed a any fractional coefficient of variation between 1 and $1/\sqrt{a}$ can be produced by a suitable choice of $\{\rho_i\}$.

The positions of the poles of Equation (6.65) determine the transition probabilities and the number of poles equals the number of stages.

When all the stages are identically distributed with parameter ρ , Expression (6.65) reduces to

$$\left(\frac{\rho}{\rho + s} \right)^a \quad (6.67)$$

and the corresponding probability density function is the Special Erlangian distribution

$$\frac{\rho(\rho x)^{a-1} e^{-\rho x}}{(a-1)!} \quad (6.68)$$

where a is a positive integer. The corresponding survivor function is

$$e^{-\rho x} \sum_{i=1}^a \frac{(\rho x)^{i-1}}{(i-1)!}$$

The characteristics of a family of Special Erlangian Distributions with a constant mean of one day are shown in Fig. 6.6. The exponential is a special case of the Special Erlangian Distribution with $\rho = 1$.

A generalization of Equation (6.68) is to replace the parameter a restricted to integer values by a parameter having any real positive value. The probability density function of Equation (6.68) then becomes the Gamma distribution

$$\frac{\rho(\rho x)^{\alpha-1} e^{-\rho x}}{\Gamma(\alpha)} \quad (6.69)$$

where the gamma function

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du$$

The mean and standard deviation of the gamma distribution are α/ρ and $\sqrt{\alpha}/\rho$ respectively and the fractional coefficient of variation is therefore $1/\sqrt{\alpha}$. For a fixed mean $\mu = \alpha/\rho$ any coefficient of variation between 1 and $1/\sqrt{\alpha}$ may be obtained by varying α and ρ in the same proportion. If μ is kept constant then as $\alpha, \rho \rightarrow \infty$, the coefficient of variation approaches zero,

184

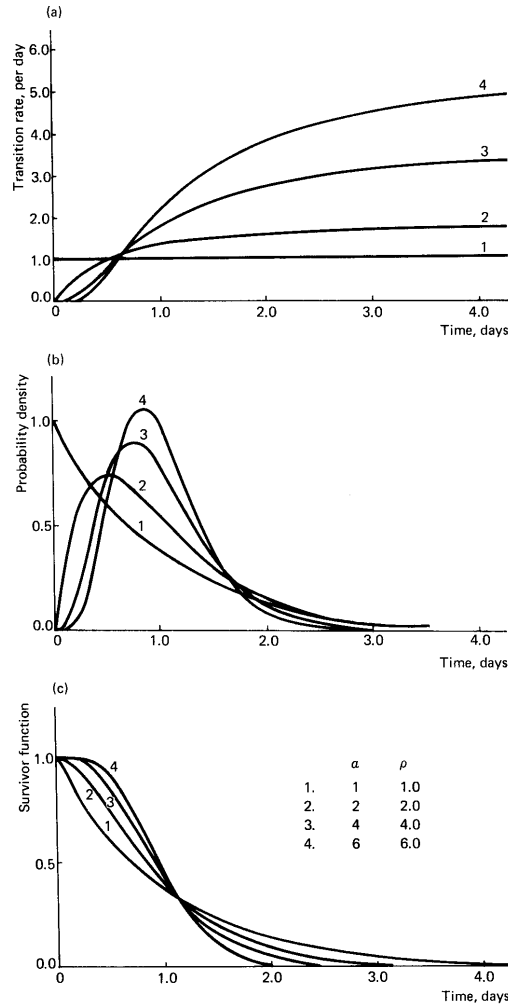


Fig. 6.6 Characteristics of the special Erlangian distribution.

i.e. there is no dispersion about the mean. This corresponds to the case of constant state duration of the component.

Many empirical distributions can be represented, at least approximately by a suitable choice of parameters α and ρ . It should, however, be realized that as α is not always an integer it may not be possible to interpret directly the distribution in terms of stages. When α is not an integer, it is preferable to solve for integer ' α ' using Equation (6.68) and then the numerical answer for α can be obtained by interpolation.

The Stages in Parallel

When the stages are in parallel, there is a probability distribution $\{\omega_i\}$, $i = 1, \dots, a$ such that the random variable X denoting the state duration of the component, has the probability ω_i of beginning on the i th stage, the life thereafter consisting of a single stage, i.e. only a single stage is used in any one realization of X . The probability density function of X is given by

$$\sum_{i=1}^a \omega_i \rho_i e^{-\rho_i x} \quad (6.70)$$

The probability density function of Equation (6.70) is formally equivalent to Equation (6.66) with the important difference that ω_i are all non-negative whereas there is no such restriction on A_i . The Laplace transform of Equation (6.70) is a rational one, i.e. the ratio of a polynomial of degree at the most ' a ' to a polynomial of the degree ' a '. If in practice, only two stages in parallel are required, the Expression (6.70) reduces to

$$\omega_1 \rho_1 e^{-\rho_1 x} + \omega_2 \rho_2 e^{-\rho_2 x} \quad (6.71)$$

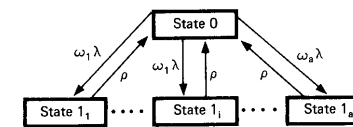


Fig. 6.7 The state transition diagram for a component having an exponentially distributed up time and whose down time has a probability density function of Equation (6.70)

It can be shown that by a suitable choice of ω_i, ρ_1, ρ_2 and the distribution of Equation (6.71) can have any desired mean and any fractional coefficient of variation between 1 and ∞ . The state transition diagram for a component whose

up time is exponentially distributed with parameter λ and whose down time has probability density function (6.70) is shown in Fig. 6.7. The total transition rate out of the up state, i.e. O is λ but it is directed towards different parallel stages of the down time according to the probability distribution $\{\omega_i\}$. That this transition diagram does represent the distribution can be seen by solving for $p_0(t)$. The differential equations are

$$p'_0(t) = -\lambda p_0(t) + \sum_{i=1}^a \rho_i p_{1i}(t) \quad (6.72)$$

$$p'_{1i}(t) = -\rho_i p_{1i}(t) + \omega_i \lambda p_0(t) \quad (6.73)$$

Taking the Laplace transform with the initial condition $p_0(0) = 1$

$$s\bar{p}_0(s) = 1 - \lambda\bar{p}_0(s) + \sum_{i=1}^a \rho_i \bar{p}_{1i}(s) \quad (6.74)$$

and

$$s\bar{p}_{1i}(s) = -\rho_i \bar{p}_{1i}(s) + \omega_i \lambda \bar{p}_0(s) \quad (6.75)$$

where $\bar{p}(s)$ is the Laplace of $p(t)$.

Substituting for $\bar{p}_{1i}(s)$ from (6.75) into (6.74) and simplifying

$$\bar{p}_0(s) = \frac{1}{D(s)} = \frac{1}{s + \lambda - \lambda \sum_i \frac{\omega_i \rho_i}{s + \rho_i}} \quad (6.76)$$

It can be easily proved that for a component having an exponentially distributed up time and having a down time distribution of $f(t)$

$$\bar{p}_0(s) = \frac{1}{s + \lambda - \lambda \bar{f}(s)} \quad (6.77)$$

where $\bar{f}(s)$ is the Laplace transform of $f(t)$.

Substituting the Laplace of the probability density function (6.70)

$$\bar{p}_0(s) = \frac{1}{s + \lambda - \lambda \sum_{i=1}^a \frac{\omega_i \rho_i}{s + \rho_i}}$$

This expression is the same as Equation (6.76) derived using the state transition diagram of Fig. 6.7. The state diagram is, therefore, an accurate representation. A generalization of Equation (6.71) is to have two series stage combinations

in parallel as shown in Fig. 6.8. The expression for the probability density function is

$$\omega_1 \rho_1 e^{-\rho_1 x} \frac{(\rho_1 x)^{a_1-1}}{(a_1-1)!} + \omega_2 \rho_2 e^{-\rho_2 x} \frac{(\rho_2 x)^{a_2-1}}{(a_2-1)!} \quad (6.78)$$

The survivor function is

$$\omega_1 e^{-\rho_1 x} \sum_{i=1}^{a_1} \frac{(\rho_1 x)^{i-1}}{(i-1)!} + \omega_2 e^{-\rho_2 x} \sum_{i=1}^{a_2} \frac{(\rho_2 x)^{i-1}}{(i-1)!}$$

The mean and the variance are

$$\text{Mean} = \omega_1 a_1 / \rho_1 + \omega_2 a_2 / \rho_2$$

$$\text{Variance} = [\omega_1 a_1 (1 + a_1) / \rho_1^2 + \omega_2 a_2 (1 + a_2) / \rho_2^2] - [\omega_1 a_1 / \rho_1 + \omega_2 a_2 / \rho_2]^2$$

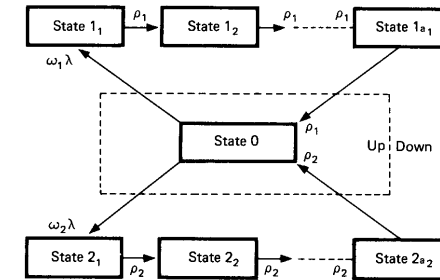


Fig. 6.8 The state transition diagram of a system with the down state represented by two series stages in parallel

The Expression (6.78) is a mixture of two Special Erlangian distributions and can approximate a wider range of distributions than the Special Erlangian. The combination has only five independent parameters. The various characteristics of this combination are shown in Fig. 6.9. These characteristics cover almost all the four types mentioned earlier. The theoretical analysis of the shape of the hazard rate function is given in Appendix II.

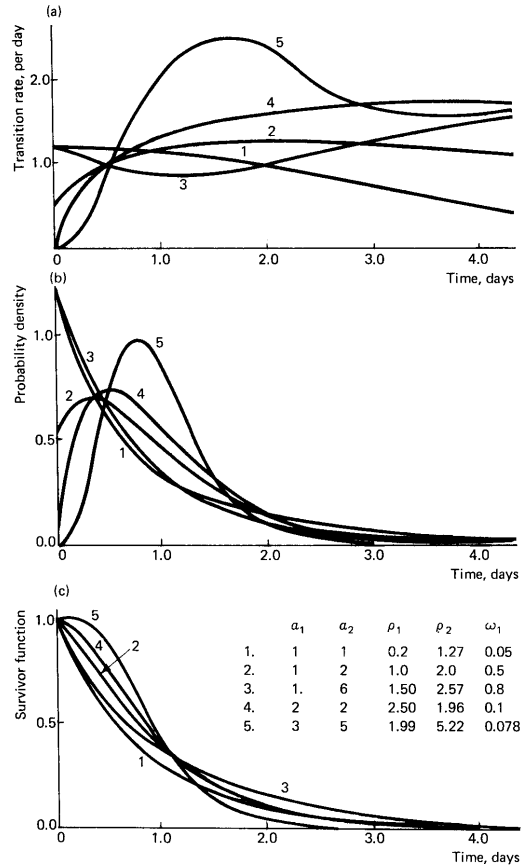


Fig. 6.9 Some characteristics of the mixture of two special Erlangian distributions

Series Stages in Series with a Distinctive Stage

The General Erlangian distribution (6.66) can generate a wider range of distribution than the Special Erlangian distribution but the number of parameters involved tends to be large. A special case combining the advantages of both is a series of identical stages in series with a distinctive stage as shown in Fig. 6.10. This model has three parameters:

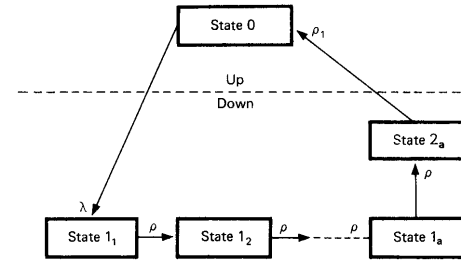


Fig. 6.10 The state transition diagram of a system with the down state represented by a number of series stages in series with a distinctive stage.

the transition rate ρ of each stage and the number of stages a in the identical series stages and the transition rate ρ_1 for the distinctive stage. The probability density function is

$$f(x) = \rho_1 \left(\frac{\rho}{\rho - \rho_1} \right)^a \left[e^{-\rho_1 x} - e^{-\rho x} \sum_{i=1}^a \frac{\{(\rho - \rho_1)x\}^{i-1}}{(i-1)!} \right] \quad (6.79)$$

The survivor function is

$$S(x) = e^{-\rho x} \sum_{i=1}^a \frac{(\rho x)^{i-1}}{(i-1)!} + \left(\frac{\rho}{\rho - \rho_1} \right)^a \times \left[e^{-\rho_1 x} - e^{-\rho x} \sum_{i=1}^a \frac{\{(\rho - \rho_1)x\}^{i-1}}{(i-1)!} \right] \quad (6.80)$$

and the hazard rate function

$$\phi(x) = \frac{f(x)}{S(x)}$$

The mean of this distribution is $a/\rho + 1/\rho_1$

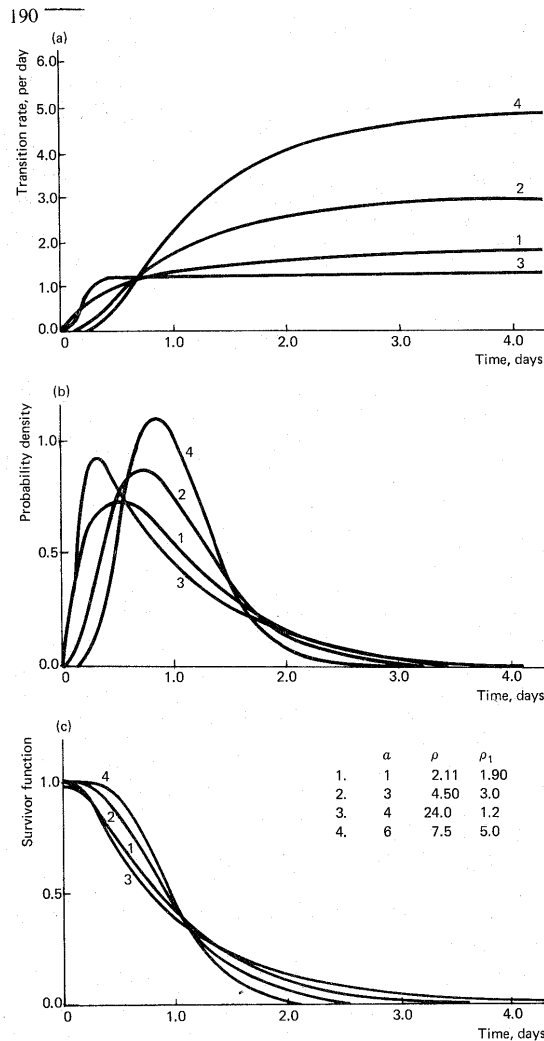


Fig. 6.11 Some characteristics of the distribution associated with series stages in series with a distinctive stage.

The characteristics of this distribution are illustrated in Fig. 6.11 and the theoretical analysis of the hazard rate is given in Appendix III.

Series Stages in Series with Two Parallel Stages

This combination has a series of identical stages followed by a stage with probability ω_1 or by another stage with probability ω_2 , as shown in Fig. 6.12a. It has five parameters and the expression for the probability density function is

$$f(x) = \omega_1 \rho_1 \left(\frac{\rho}{\rho - \rho_1} \right)^a \left[e^{-\rho_1 x} - e^{-\rho x} \sum_{i=1}^a \frac{\{(\rho - \rho_1)x\}^{i-1}}{(i-1)!} \right] + \omega_2 \rho_2 \left(\frac{\rho}{\rho - \rho_2} \right)^a \left[e^{-\rho_2 x} - e^{-\rho x} \sum_{i=1}^a \frac{\{(\rho - \rho_2)x\}^{i-1}}{(i-1)!} \right] \quad (6.81)$$

The survivor function can be expressed by

$$S(x) = \sum_{i=1}^a \frac{(\rho x)^{i-1}}{(i-1)!} e^{-\rho x} + \omega_1 \left(\frac{\rho}{\rho - \rho_1} \right)^a \times \left[e^{-\rho_1 x} - e^{-\rho x} \sum_{i=1}^a \frac{\{(\rho - \rho_1)x\}^{i-1}}{(i-1)!} \right] + \omega_2 \left(\frac{\rho}{\rho - \rho_2} \right)^a \left[e^{-\rho_2 x} - e^{-\rho x} \sum_{i=1}^a \frac{\{(\rho - \rho_2)x\}^{i-1}}{(i-1)!} \right] \quad (6.82)$$

The transition rate function can be found by

$$\phi(x) = f(x)/S(x)$$

The derivation of these expressions and the theoretical analysis of the hazard rate function is provided in Appendix IV. Comparing Expressions (6.79) and (6.81) this combination is equivalent to two 'series stages in series with a distinctive stage' in parallel as shown in Fig. 6.12b. The various characteristics of this combination are illustrated in Fig. 6.13.

Determination of Parameters

After a model is chosen to approximate a distribution, the next problem is to find the model parameters to fit the distribution. There are generally no explicit formulae for directly deriving the approximate stage model parameters from those of the distribution. In many cases, the parameters that will best define an

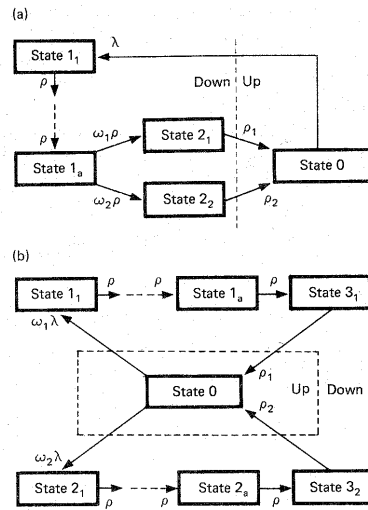


Fig. 6.12 The state transition diagram of a system with the down state represented by:
 (a) A series stage in series with two parallel stages
 (b) Two 'series stages in series with a distinctive stage' in parallel

empirical distribution are not known. The moments can, however, always be evaluated for any distribution either by exact or approximate methods. A method of determining the parameters for approximate stage models by matching the first r moments of the model and the distribution is presented in the following sections. This method is quite general in application.

The parameters of the stage model are non-linear and implicit functions of its moments. On the contrary, the first r moments for the stage combinations discussed in this paper can be easily calculated from the parameters. The Newton–Raphson method of successive approximation is applied to solve for the parameters from the given moments. This method requires for each stage of approximation:

1. evaluation of the moments with the given parameters
2. evaluation of the partial derivatives of the moments with respect to each parameter

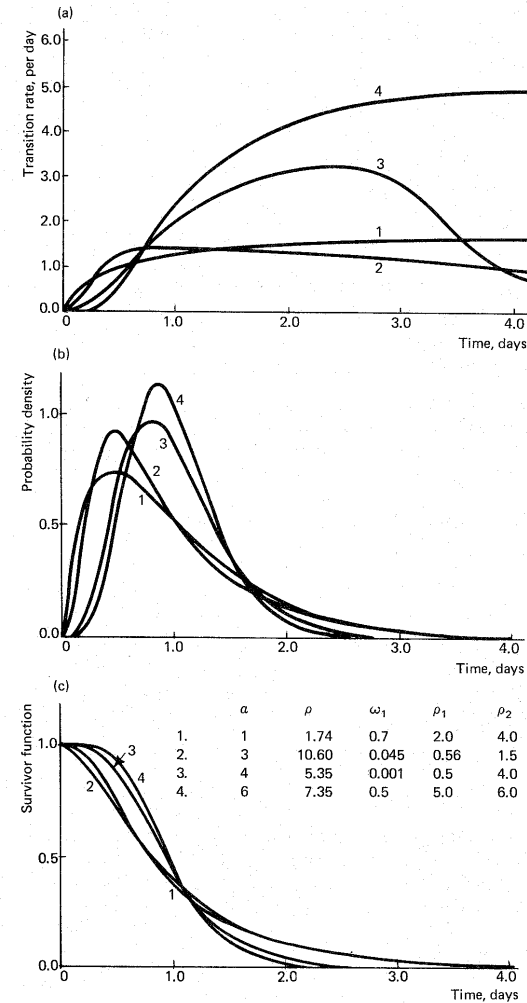


Fig. 6.13 Some characteristics of the distribution associated with series stages in series with two parallel stages.

1 Moment Evaluation for a Combination of Stages

The probability density functions of the stage models discussed have simple rational Laplace transforms. The r th moment about zero can be obtained if the r th derivative of the Laplace of the probability density function exists at $s = 0$. The r th moment m_r of the distribution is

$$m_r = (-1)^r \bar{f}^{(r)}(0)$$

where

$$\bar{f}^{(r)}(0) = \left. \frac{d^r \bar{f}(s)}{ds^r} \right|_{s=0}$$

$\bar{f}(s)$ being the Laplace transform of the probability density function. Moment calculations for some of the stage models are given in Appendix V.

2. Newton Raphson Method for Parameter Calculation

The first r moments are matched by successive approximation starting from the initial parameters. If a model has r parameters, $x_1, x_2, \dots, x_r$ to be determined by matching the first r moments, the r functions, $\phi_1, \phi_2, \dots, \phi_r$ are defined such that

$$\phi_1 = \phi_1(X) = m_1(X) - M_1$$

$$\phi_2 = \phi_2(X) = m_2(X) - M_2$$

...

$$\phi_r = \phi_r(X) = m_r(X) - M_r$$

where X is the column vector $(x_1, x_2, \dots, x_r)^t$ and $m_r(X)$ is the r th moment of the stage model and M_r is the r th moment of the distribution to be approximated. The conditions of exact match of the first r moments are

$$\phi_1 = \phi_2 = \dots = \phi_r = 0$$

Let $X_0 = (x_{10}, x_{20}, \dots, x_{r0})^t$ be the vector of the parameters at a certain stage of approximation, and let ϕ be a column vector such that $\phi = (\phi_1, \phi_2, \dots, \phi_r)^t$. The correction vector for parameter $\Delta X = (\Delta x_1, \Delta x_2, \dots, \Delta x_r)^t$, can be calculated from the following matrix equation by the Gauss elimination method if $\phi(X_0)$ and $\phi'(X_0)$ are known.

$$\phi(X_0) + \phi'(X_0)\Delta X = 0$$

where $\phi(X_0)$ is the vector ϕ and $X = X_0$, $\phi'(X_0)$ is the Jacobian matrix of ϕ at X_0 , i.e.

$$\phi'(X_0) = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1}(X_0) & \frac{\partial \phi_1}{\partial x_2}(X_0) & \dots & \frac{\partial \phi_1}{\partial x_r}(X_0) \\ \frac{\partial \phi_2}{\partial x_1}(X_0) & \frac{\partial \phi_2}{\partial x_2}(X_0) & \dots & \frac{\partial \phi_2}{\partial x_r}(X_0) \\ \dots & \dots & \dots & \dots \\ \frac{\partial \phi_r}{\partial x_1}(X_0) & \frac{\partial \phi_r}{\partial x_2}(X_0) & \dots & \frac{\partial \phi_r}{\partial x_r}(X_0) \end{bmatrix}$$

The improved parameter values are obtained by $X = X_0 + \Delta X$. The $\phi(X_0)$ can be calculated directly from the first r moments of the model when $X = X_0$ (Appendix VI).

The method for finding $\phi'(X_0)$ is discussed for some of the stage models in Appendix VI.

Example system study

The technique of stages is applied to a two state unit having the up time exponentially distributed and down time with a lognormal distribution. The following numerical values are used

$$\text{Mean Up Time} = 1500 \text{ hr}$$

$$\text{Mean Down Time} = 20 \text{ hr}$$

The standard deviation of the down time is varied as

$$(i) \quad 10 \text{ hr}$$

$$(ii) \quad 14.14 \text{ hr}$$

$$(iii) \quad 20 \text{ hr}$$

The lognormal distribution is completely specified by its mean and standard deviation. The expression for a lognormal probability density function of the random variable X was given in Chapter 2.

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\log x - m)^2 / 2\sigma^2}$$

where m and σ are the standard deviation of $\log(X)$. The r th moment of X is

$$m_r(X) = E(X^r) = e^{mr + \sigma^2 r^2 / 2} \quad (6.83)$$

The mean is

$$m_X = e^{m + \sigma^2 / 2} \quad (6.84)$$

and the variance is

$$\sigma_X^2 = e^{2m + 2\sigma^2} - e^{2m + \sigma^2} \quad (6.85)$$

Solving Equations (6.84) and (6.85)

$$\sigma^2 = \log \left[\left(\frac{\sigma_X}{m_X} \right)^2 + 1 \right]$$

and

$$m = \log m_X - \frac{1}{2} \log \left[\left(\frac{\sigma_X}{m_X} \right)^2 + 1 \right]$$

The parameters m and σ can therefore be found from the mean m and standard deviation σ_X of the log normal distribution. The hazard rate of a lognormal distribution shows initial positive ageing followed by negative ageing.

This suggests two combinations:

- two series stages in parallel
- series stages in series with two parallel stages

The log normal is approximated by these two combinations using moment matching technique and the parameters are listed in Table 6.1. The transition rate functions, probability density functions and the survivor functions of the log normal distributions together with the stage combination models are shown in Fig. 6.14

Table 6.1 The parameters of stage combinations for the approximation of lognormal distributions

The mean of the distribution $m_X = 20$ hours, $\sigma_X =$ Standard deviation of the distribution

(a) Approximation by series stages in series with two parallel stages

Lognormal σ_X (hours)	Approximate model parameters				
	a	ω_1	ρ per hour	ρ_1 per hour	ρ_2 per hour
10	6	0.10976	0.526715	0.078325	0.123519
14.14	4	0.05270	0.554433	0.036597	0.083496
20	3	0.02250	1.221384	0.016180	0.060515

(b) Approximation by two series stages in parallel

Lognormal σ_X (hours)	Approximate model parameters				
	a_1	a_2	ω_1	ρ_1 per hour	ρ_2 per hour
10	4	6	0.21601	0.144001	0.336001
14.14	2	4	0.25464	0.066666	0.241201
20	2	2	0.06699	0.031698	0.118301

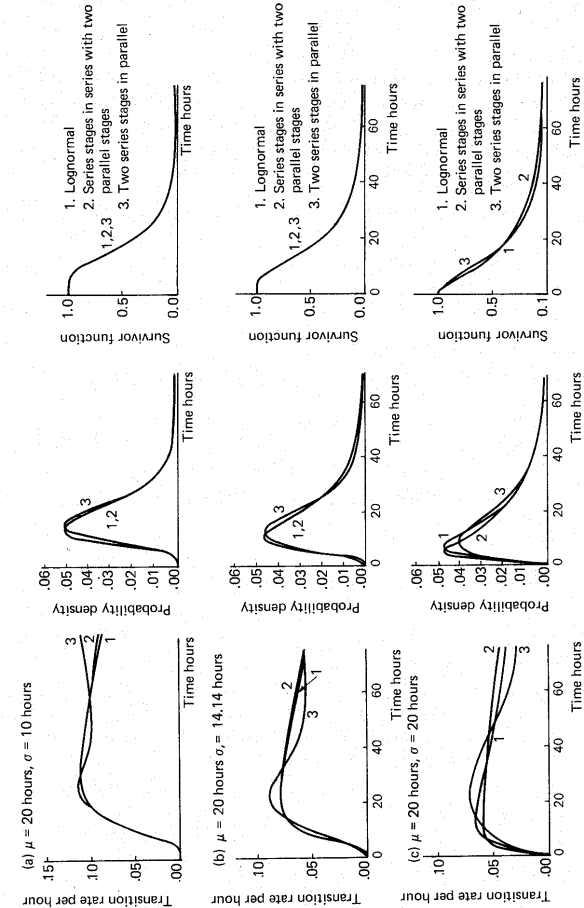


Fig. 6.14 Characteristics of lognormal distributions and their approximate stage combination models.

Several techniques exist for finding time specific probabilities of the states of Markovian process. The probability for the down state was evaluated up to 24 hours assuming the system was initially in the up state. The continuous time Markov process is approximated by a discrete time process using a small time interval and the state probabilities are obtained by multiplication of the transition probability matrix. The results are shown in Table 6.2. The results compare well with the direct approximate expression derived by assuming that the time period considered is so short that no more than one forced outage and repair can occur in it. Denoting the up and down states by 1 and 2 respectively, and assuming $p_1(0) = 1$

$$p_2(t) \approx p_2^a(t) = P(\text{one forced outage and no repair in } t) \\ = \int_0^t \lambda e^{-\lambda x} S_d(t-x) dx$$

where

λ = failure rate

S_d = survivor function of down time

and

$p_2^a(t)$ = the approximate expression for the probability of being in the down state at time t

Assuming the interval $(0, t)$ to be divided into k equal subintervals each of length δ

$$p_2^a(t) = \sum_{i=1}^k [P(\text{forced outage in } i\text{th interval})P(\text{no repair in interval of length } (k-i)\delta)] \\ = \sum_{i=1}^k [e^{-\lambda(i-1)\delta} - e^{-\lambda i\delta}] S_d\{(k-i)\delta\}$$

when $\lambda t \ll 1$

$$e^{-\lambda t} = 1 - \lambda t$$

Therefore

$$p_2^a(t) = \lambda \delta \sum_{i=1}^k S_d\{(k-i)\delta\} \\ = \lambda \delta \sum_{i=1}^k S_d\{(i-1)\delta\} \\ = \lambda \delta \sum_{i=0}^{k-1} S_d\{i\delta\}$$

It may be necessary to find the state probability values at the end of each equal time interval, i.e. in a 24 hr period, calculate $p_2^a(t)$ at the end of each

hour. The probability at the end of the n th interval is given by

$$p_2^a(nm\delta) = p_2^a\{(n-1)m\delta\} + \lambda \delta \sum_{i=(n-1)m}^{nm-1} S_d(i\delta)$$

where m is the number of subintervals in each time interval and δ is the length of each subinterval, the length of each interval therefore being $m\delta$. The closed form expression for the survivor function of the lognormal distribution does not exist and Simpson's Rule can be applied to calculate this by integrating the probability density function. The following equation can be used to calculate $S_d(i\delta)$ successively

$$S_d(i\delta) = S\{(i-1)\delta\} - \frac{\delta}{6l} \sum_{j=0}^{l-1} \left[f\left\{(i-1)\delta + \frac{j\delta}{l}\right\} \right. \\ \left. + 4f\left\{(i-1)\delta + \left(j + \frac{1}{2}\right)\frac{\delta}{l}\right\} + f\left\{(i-1)\delta + (j+1)\frac{\delta}{l}\right\} \right]$$

where l is the number of divisions in each subinterval for calculation of the survivor function by Simpson's Rule. This method was developed for computer application without requiring excessive storage. Down state probabilities are evaluated up to 24 hours with a one hour interval. They are listed in Table 6.2 together with those obtained by the device of stages. It can be seen that the results of the two approximate methods are quite close to each other.

Application to The Transformer Bank Problem

Reliability Modelling for Three Single Phase Transformers With One Spare, General distributions for Repair and Change Out, Restricted Repair, Using Stages in Parallel.

The expressions for this model assuming general distributions have already been developed. However, to illustrate the modelling approach when the stages are in parallel, the state transition diagram is developed in Fig. 6.15. Since the change out operation is restricted to a single stage, from the point of view of steady state analysis, the change out rate can be simply represented by an average value γ .

The probability density function for the repair time is

$$\omega_1 \mu_1 e^{-\mu_1 x} + \omega_2 \mu_2 e^{-\mu_2 x} \quad \text{with mean} = \omega_1/\mu_1 + \omega_2/\mu_2 \quad (6.86)$$

The state transition diagram in Fig. 6.15 can be clarified by referring to the discussion related to Fig. 6.7. The first letter of the state numbers represents

Table 6.2 Time specific probabilities of the down state calculated by using

(a) The approximate expression (b) The approximate 'series stages in series with two parallel stages' (c) the approximate 'two series stages in parallel'. The Mean Down Time = 20 hours. The Mean Up Time = 1500 hours

Time (hours)	Std. deviation = 10 hours			Std. deviation = 14.4 hours			Std. deviation = 20 hours		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
1	6.667 × 10 ⁻⁴	6.667 × 10 ⁻⁴	6.667 × 10 ⁻⁴	6.667 × 10 ⁻⁴	6.674 × 10 ⁻⁴	6.670 × 10 ⁻⁴	6.666 × 10 ⁻⁴	6.680 × 10 ⁻⁴	6.654 × 10 ⁻⁴
2	1.333 × 10 ⁻³	1.334 × 10 ⁻³	1.334 × 10 ⁻³	1.333 × 10 ⁻³	1.334 × 10 ⁻³	1.332 × 10 ⁻³	1.331 × 10 ⁻³	1.330 × 10 ⁻³	1.323 × 10 ⁻³
3	2.000 × 10 ⁻³	2.000 × 10 ⁻³	2.000 × 10 ⁻³	2.000 × 10 ⁻³	1.998 × 10 ⁻³	1.995 × 10 ⁻³	1.985 × 10 ⁻³	1.974 × 10 ⁻³	1.966 × 10 ⁻³
4	2.008 × 10 ⁻³	2.686 × 10 ⁻³	2.666 × 10 ⁻³	2.000 × 10 ⁻³	2.658 × 10 ⁻³	2.651 × 10 ⁻³	2.620 × 10 ⁻³	2.591 × 10 ⁻³	2.590 × 10 ⁻³
5	3.800 × 10 ⁻³	3.338 × 10 ⁻³	3.328 × 10 ⁻³	3.312 × 10 ⁻³	3.308 × 10 ⁻³	3.297 × 10 ⁻³	3.280 × 10 ⁻³	3.178 × 10 ⁻³	3.192 × 10 ⁻³
6	3.000 × 10 ⁻³	3.988 × 10 ⁻³	3.986 × 10 ⁻³	3.950 × 10 ⁻³	3.943 × 10 ⁻³	3.930 × 10 ⁻³	3.811 × 10 ⁻³	3.732 × 10 ⁻³	3.770 × 10 ⁻³
7	4.650 × 10 ⁻³	4.640 × 10 ⁻³	4.636 × 10 ⁻³	4.567 × 10 ⁻³	4.557 × 10 ⁻³	4.545 × 10 ⁻³	4.362 × 10 ⁻³	4.255 × 10 ⁻³	4.321 × 10 ⁻³
8	5.800 × 10 ⁻³	5.280 × 10 ⁻³	5.278 × 10 ⁻³	5.160 × 10 ⁻³	5.147 × 10 ⁻³	5.139 × 10 ⁻³	4.880 × 10 ⁻³	4.748 × 10 ⁻³	4.846 × 10 ⁻³
9	5.928 × 10 ⁻³	5.983 × 10 ⁻³	5.895 × 10 ⁻³	5.725 × 10 ⁻³	5.709 × 10 ⁻³	5.708 × 10 ⁻³	5.367 × 10 ⁻³	5.212 × 10 ⁻³	5.342 × 10 ⁻³
10	6.560 × 10 ⁻³	6.595 × 10 ⁻³	6.496 × 10 ⁻³	6.268 × 10 ⁻³	6.240 × 10 ⁻³	6.249 × 10 ⁻³	5.823 × 10 ⁻³	5.650 × 10 ⁻³	5.812 × 10 ⁻³
11	7.110 × 10 ⁻³	7.081 × 10 ⁻³	7.072 × 10 ⁻³	6.765 × 10 ⁻³	6.740 × 10 ⁻³	6.761 × 10 ⁻³	6.250 × 10 ⁻³	6.061 × 10 ⁻³	6.254 × 10 ⁻³
12	7.000 × 10 ⁻³	7.628 × 10 ⁻³	7.620 × 10 ⁻³	7.288 × 10 ⁻³	7.208 × 10 ⁻³	7.243 × 10 ⁻³	6.649 × 10 ⁻³	6.449 × 10 ⁻³	6.669 × 10 ⁻³
13	8.178 × 10 ⁻³	8.143 × 10 ⁻³	8.137 × 10 ⁻³	7.680 × 10 ⁻³	7.644 × 10 ⁻³	7.694 × 10 ⁻³	7.022 × 10 ⁻³	6.815 × 10 ⁻³	7.059 × 10 ⁻³
14	8.681 × 10 ⁻³	8.626 × 10 ⁻³	8.021 × 10 ⁻³	8.092 × 10 ⁻³	8.050 × 10 ⁻³	8.114 × 10 ⁻³	7.370 × 10 ⁻³	7.159 × 10 ⁻³	7.424 × 10 ⁻³
15	9.109 × 10 ⁻³	9.068 × 10 ⁻³	9.070 × 10 ⁻³	8.475 × 10 ⁻³	8.427 × 10 ⁻³	8.503 × 10 ⁻³	7.696 × 10 ⁻³	7.483 × 10 ⁻³	7.765 × 10 ⁻³
16	9.522 × 10 ⁻³	9.477 × 10 ⁻³	9.488 × 10 ⁻³	8.830 × 10 ⁻³	8.776 × 10 ⁻³	8.862 × 10 ⁻³	8.000 × 10 ⁻³	7.789 × 10 ⁻³	8.088 × 10 ⁻³
17	9.901 × 10 ⁻³	9.852 × 10 ⁻³	9.863 × 10 ⁻³	9.160 × 10 ⁻³	9.099 × 10 ⁻³	9.193 × 10 ⁻³	8.285 × 10 ⁻³	8.077 × 10 ⁻³	8.380 × 10 ⁻³
18	1.025 × 10 ⁻²	1.019 × 10 ⁻²	1.021 × 10 ⁻²	9.464 × 10 ⁻³	9.398 × 10 ⁻³	9.498 × 10 ⁻³	8.551 × 10 ⁻³	8.318 × 10 ⁻³	8.657 × 10 ⁻³
19	1.056 × 10 ⁻²	1.050 × 10 ⁻²	1.052 × 10 ⁻²	9.746 × 10 ⁻³	9.674 × 10 ⁻³	9.776 × 10 ⁻³	8.800 × 10 ⁻³	8.604 × 10 ⁻³	8.916 × 10 ⁻³
20	1.085 × 10 ⁻²	1.078 × 10 ⁻²	1.080 × 10 ⁻²	1.001 × 10 ⁻²	9.993 × 10 ⁻³	1.008 × 10 ⁻²	9.084 × 10 ⁻³	8.845 × 10 ⁻³	9.155 × 10 ⁻³
21	1.111 × 10 ⁻²	1.104 × 10 ⁻²	1.106 × 10 ⁻²	1.025 × 10 ⁻²	1.017 × 10 ⁻²	1.026 × 10 ⁻²	9.253 × 10 ⁻³	9.072 × 10 ⁻³	9.878 × 10 ⁻³
22	1.134 × 10 ⁻²	1.126 × 10 ⁻²	1.129 × 10 ⁻²	1.047 × 10 ⁻²	1.038 × 10 ⁻²	1.048 × 10 ⁻²	9.458 × 10 ⁻³	9.286 × 10 ⁻³	9.586 × 10 ⁻³
23	1.155 × 10 ⁻²	1.146 × 10 ⁻²	1.149 × 10 ⁻²	1.067 × 10 ⁻²	1.058 × 10 ⁻²	1.067 × 10 ⁻²	9.650 × 10 ⁻³	9.486 × 10 ⁻³	9.778 × 10 ⁻³
24	1.174 × 10 ⁻²	1.165 × 10 ⁻²	1.167 × 10 ⁻²	1.085 × 10 ⁻²	1.077 × 10 ⁻²	1.085 × 10 ⁻²	9.881 × 10 ⁻³	9.678 × 10 ⁻³	9.958 × 10 ⁻³

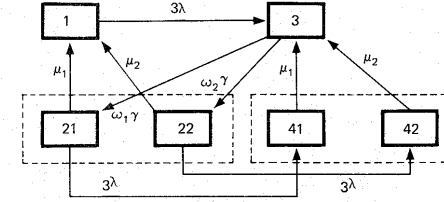


Fig. 6.15 The state transition diagram for three single phase transformers with one spare, assuming restricted repair and two stages in parallel for the repair time.

the system state as given in Fig. 6.2, and the second letter if any denotes the stage number. The steady state equations for this model can be written in the conventional manner.

The solution of these equations gives

$$p_1 = 1/Z \quad \text{where} \quad Z = 1 + \frac{3\lambda}{E} \left(\frac{1}{\gamma} + \frac{1}{\mu} \right)$$

where

$$E = \frac{\omega_1 \mu_1}{3\lambda + \mu_1} + \frac{\omega_2 \mu_2}{3\lambda + \mu_2} \quad \text{and} \quad \frac{1}{\mu} = \frac{\omega_1}{\mu_1} + \frac{\omega_2}{\mu_2}$$

also

$$p_2 = p_{21} + p_{22} = \frac{1-E}{ZE}, \quad p_3 = \frac{3\lambda}{\gamma EZ} \quad \text{and} \quad p_4 = p_{41} + p_{42} = \frac{9\gamma^2}{ZE} \left(\frac{\omega_1}{\mu_1(3\lambda + \mu_1)} + \frac{\omega_2}{\mu_2(3\lambda + \mu_2)} \right)$$

Three Single Phase Transformers With a Single Spare, Unrestricted Repair, Erlangian Distribution for Repair and Reinstallation.

The state transition diagram shown in Fig. 6.16, assumes three stages for each of the repair and reinstallation periods, i.e. the index 'a' in Equation (6.68) is 3. The transition diagrams for higher 'a' can be similarly developed and the numerical answers when a is a non-integer, (i.e. when Equation (6.69) is assumed as the distribution) can be obtained by interpolation. A computer program can be written which can generate the transition rate matrix for any 'a' without having to draw the state transition diagram. The reliability evaluation can, therefore, be made conveniently for any value of 'a'.

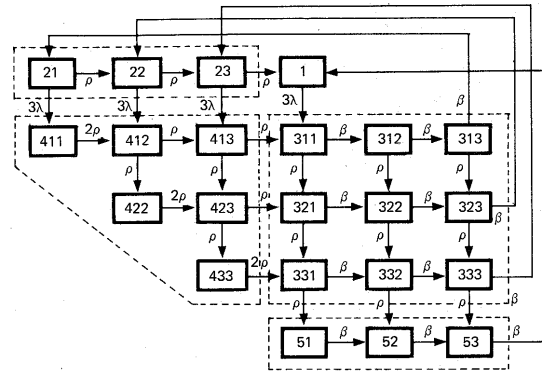


Fig. 6.16 The state transition diagram for three single phase transformers with a single spare assuming unrestricted repair and a general Erlangian distribution for the repair and reinstallation period.

The repair and the change out rates associated with each stage are assumed as ρ and β respectively such that $\rho = a \cdot \mu$ and $\beta = a \cdot \gamma$. The first letter of the state number in Fig. 6.16 indicates the system state as shown in Fig. 6.1 and the remaining letters denote the stages of the operations being carried out. For example, 22 means that the system is in state 2 (the transformer bank is up and no spare) and the repair is in the second stage. Similarly, 412 indicates state 4 (transformer bank failed and no spare) and the repair on one of the transformers is in the first stage and the other transformer in the second stage. As another example, 323 denotes state 3 (transformer bank down and one spare) and the repair is in the second stage whereas the change out operation is in stage 3. The relationship between the state and stage availabilities is

$$p_1 = p_1, \quad p_2 = \sum_{i=1}^a p_{2i}, \quad p_3 = \sum_{i=1}^a \sum_{j=1}^a p_{3ij},$$

$$p_4 = \sum_{i=1}^a \sum_{j=1}^a p_{4ij}, \quad p_5 = \sum_{i=1}^a p_{5i}$$

The steady state equations for Fig. 6.16 can be written in the form

$$AP = B \quad (6.87)$$

where

$B = A$ column vector with all zeros

$P =$ the vector of steady state probabilities

and

$A =$ the transpose of the transition rate matrix

Any $(a-1)$ equations out of a in (6.87) can be solved with the normalizing equation

$$\sum_{i=1}^5 p_i = 1$$

to give the steady state probabilities.

Numerical Results

These studies were conducted to determine the effect of the distribution form on the steady state probabilities. The exponential distribution was assumed for the up time and the repair and change out periods were assumed to have the Special Erlangian distribution of Equation (6.68) with the same shape parameter ' a '.

Restricted Repair

The failure rate was taken as 0.008 failures per year and the effect of variation of ' a ' on the probability of being in the down state was determined under different values for the mean repair and change out times using Equation (6.30). The results are shown in Tables 6.3 and 6.4. It should be noted that the mean values are held constant so that the difference in values is entirely due to the change in the shape of the distribution. The values with $a = 1$ correspond to the exponential distribution and the limiting values (LV) refer to the constant repair and reinstallation times i.e. when $a \rightarrow \infty$ with the mean values kept constant.

Table 6.3 shows the actual values of the steady state unavailability and Table 6.4 shows the percentage variation from the exponential when the value of ' a ' is increased. It can be seen that the value of ' a ' has a considerable effect on the unavailability of the transformer bank. The variation from the exponential depends both upon the mean repair time (MRT) and the mean change out time (MIT). For a given mean change out time, the greater the mean repair time the more pronounced is the variation from the exponential and for a given mean repair time, the greater the mean change out time the less pronounced the variation.

Table 6.3 The effect of a on the unavailability of the transformer bank

a	M.R.T. = 182.5 Days		M.R.T. = 20 Days	
	M.I.T. = 0.5 (Days)	M.I.T. = 3.5 (Days)	M.I.T. = 0.5 (Days)	M.I.T. = 3.5 (Days)
1	0.175139×10^{-3}	0.372291×10^{-3}	0.346026×10^{-4}	0.231810×10^{-3}
2	0.140000×10^{-3}	0.337166×10^{-3}	0.341714×10^{-4}	0.231379×10^{-3}
3	0.128224×10^{-3}	0.325395×10^{-3}	0.340276×10^{-4}	0.231236×10^{-3}
4	0.122325×10^{-3}	0.319498×10^{-3}	0.339557×10^{-4}	0.231164×10^{-3}
5	0.118781×10^{-3}	0.315956×10^{-3}	0.339126×10^{-4}	0.231121×10^{-3}
LV	0.104546×10^{-3}	0.301726×10^{-3}	0.337212×10^{-4}	0.230929×10^{-3}

Table 6.4 The percent variation from the values of exponential

a	M.R.T. = 182.5 Days		M.R.T. = 20 Days	
	M.I.T. = 0.5 (Days)	M.I.T. = 3.5 (Days)	M.I.T. = 0.5 (Days)	M.I.T. = 3.5 (Days)
1	0.00	0.00	0.00	0.00
2	20.06	9.43	1.25	0.18
3	26.79	12.60	1.66	0.25
4	30.16	14.18	1.87	0.28
5	32.18	15.13	1.99	0.30
LV	40.31	18.95	2.55	0.38

Unrestricted Repair

A similar study was conducted for unrestricted repair using the computer program which generates the transition rate matrix for different values of ' a ' and evaluates the various steady state probabilities. The variation of the unavailability with the increase in ' a ' is shown in Table 6.5A and similar variations in the states constituting the failure state, i.e. p_3 , p_4 and p_5 are shown in Tables 6.5B – D. The probability density functions for both repair and change out are assumed to be Special Erlangian.

The exponential distribution, as in the case of restricted repair, does overestimate the unavailability but the variation is insignificant. The individual components of the failure state show an interesting behaviour; p_3 shows an increase with increasing ' a ' whereas p_4 decreases. The only component which shows a large variation is p_5 , however, its magnitude is relatively quite small. The large variation in the value of p_5 is explained by the fact that an increase in ' a ' is accompanied by a rapid decrease of dispersion of the repair and change

out times which results in a lower probability of being repaired while the change out is in progress. The overall variation in the unavailability is, however, quite small.

Two Three Phase Transformers in Parallel, Special Erlangian Distribution for Repair and Change Out Periods

The combination of stages to approximate the probability density function depends upon the available information. However, a model has been developed

Table 6.5 The variation in the probability of the failure state and its constituent states

a	M.R.T. = 182.5 Days		M.R.T. = 20 Days	
	M.I.T. = 0.5 (Days)	M.I.T. = 3.5 (Days)	M.I.T. = 0.5 (Days)	M.I.T. = 3.5 (Days)
A. Variation in Unavailability				
1	0.103818×10^{-3}	0.299854×10^{-3}	0.337181×10^{-4}	0.230819×10^{-3}
2	0.103769×10^{-3}	0.299513×10^{-3}	0.337126×10^{-4}	0.230780×10^{-3}
3	0.103745×10^{-3}	0.299344×10^{-3}	0.337099×10^{-4}	0.230760×10^{-3}
4	0.103731×10^{-3}	0.299237×10^{-3}	0.337081×10^{-4}	0.230748×10^{-3}
5	0.103720×10^{-3}	0.299164×10^{-3}	0.337068×10^{-4}	0.230739×10^{-3}
B. Variation in p_3				
1	0.327835×10^{-4}	0.225739×10^{-3}	0.320738×10^{-4}	0.195816×10^{-3}
2	0.328720×10^{-4}	0.229870×10^{-3}	0.328363×10^{-4}	0.220615×10^{-3}
3	0.328726×10^{-4}	0.230027×10^{-3}	0.328726×10^{-4}	0.226905×10^{-3}
4	0.328726×10^{-4}	0.230037×10^{-3}	0.328748×10^{-4}	0.228900×10^{-3}
5	0.328726×10^{-4}	0.230039×10^{-3}	0.328749×10^{-4}	0.229609×10^{-3}
C. Variation in p_4				
1	0.709445×10^{-4}	0.697866×10^{-4}	0.842477×10^{-6}	0.734782×10^{-6}
2	0.708962×10^{-4}	0.694447×10^{-4}	0.836972×10^{-6}	0.695249×10^{-6}
3	0.708722×10^{-4}	0.692747×10^{-4}	0.834234×10^{-6}	0.674866×10^{-6}
4	0.708574×10^{-4}	0.691686×10^{-4}	0.832526×10^{-6}	0.662111×10^{-6}
5	0.708467×10^{-4}	0.690945×10^{-4}	0.831331×10^{-6}	0.653233×10^{-6}
D. Variation in p_5				
1	0.898177×10^{-7}	0.432924×10^{-5}	0.801844×10^{-6}	0.342618×10^{-4}
2	0.128310×10^{-8}	0.197985×10^{-6}	0.393765×10^{-7}	0.846884×10^{-5}
3	0.713079×10^{-9}	0.415356×10^{-7}	0.301219×10^{-8}	0.318000×10^{-5}
4	0.665219×10^{-9}	0.323045×10^{-7}	0.808461×10^{-9}	0.118599×10^{-5}
5	0.638378×10^{-9}	0.305947×10^{-7}	0.636642×10^{-9}	0.477498×10^{-5}

in Fig. 6.17, assuming the Special Erlangian distribution for repair and change out because by suitably varying the parameter ' a ' the behaviour of a large number of probability density functions less dispersed than the exponential can be approximated. The same symbols and notation as in Fig. 6.16 has been used. The transition diagram shown is for three stages. A computer program can be

easily written to generate the transition rate matrix for any value of 'a'. The numerical values for time dependent or steady state probabilities can then be obtained.

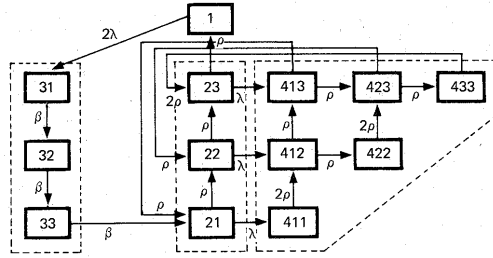


Fig. 6.17 The state transition diagram for two three phase half capacity transformers in parallel assuming the special Erlangian distribution for repair and change out times.

The Special Case of Independent Components

If a system is composed of independent binary components, the associated stochastic process is a superposition of independent alternating renewal processes. It was proved in Chapter 3 for an equilibrium alternating renewal process that irrespective of the probability density function of the up and down time durations, the frequency of encountering the up and down states is given by

$$f_u = f_d = \frac{1}{T_u + T_d}$$

It is also known that for an equilibrium alternating renewal process, the probabilities of being in the up and down state respectively are given by

$$p_u = \frac{T_u}{T_u + T_d}$$

and

$$p_d = \frac{T_d}{T_u + T_d}$$

The transition rates from up to down, λ , and down to up, μ , are therefore, given by

$$\lambda = f_d/p_u = \frac{1}{T_u} \quad \text{and} \quad \mu = f_u/p_d = \frac{1}{T_d}$$

It is therefore clear that under steady state conditions, the interstate transition rates of an alternating renewal process can be represented by the reciprocals of the respective mean state durations. Since the different alternating renewal processes are independent, the steady state probabilities and frequencies will be unaffected by the forms of the probability density functions of the state durations provided the transition rates are represented by the mean component state durations.

Reliability Modelling Using Complex Transition Rates

It has been noted that when complex transition rates are allowed, any distribution having a rational Laplace transform can, in principle, be treated using the state device. The use of complex transition rates is illustrated for a component whose up time is exponentially distributed with rate parameter λ and down time has the density function

$$f(x) = a \frac{(a^2 + b^2)}{b^2} e^{-at} (1 - \cos bt) \quad (6.88)$$

The Laplace transform is

$$\begin{aligned} \bar{f}(s) &= \frac{a}{a+s} \frac{a^2 + b^2}{(a+s)^2 + b^2} \\ &= \frac{a}{a+s} \frac{a+ib}{(a+ib)+s} \frac{a-ib}{(a-ib)+s} \end{aligned}$$

This expression is the product of the Laplace of three functions as shown below

Laplace	function
$\frac{a}{a+s}$	$a e^{-at}$
$\frac{a+ib}{(a+ib)+s}$	$(a+ib) e^{-(a+ib)t}$
$\frac{a-ib}{(a-ib)+s}$	$(a-ib) e^{-(a-ib)t}$

Equation (6.88), therefore, is the probability density function of the random variable which is the sum of three random variables having exponential distributions. The state transition diagram of this component is, therefore, as

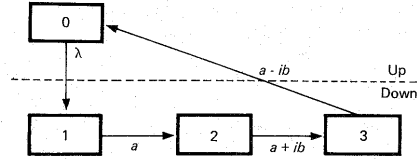


Fig. 6.18 State transition diagram of a component having an exponential up time distribution and a down time distribution given by Equation (6.88)

shown in Fig. 6.18. The state differential equations can be written as below

$$p'_0(t) = -\lambda p_0(t) + (a - ib)p_3(t)$$

$$p'_1(t) = -ap_1(t) + \lambda p_0(t)$$

$$p'_2(t) = -(a + ib)p_2(t) + ap_1(t)$$

$$p'_3(t) = -(a - ib)p_3(t) + (a + ib)p_2(t)$$

Assuming $p_0(0) = 1$, the Laplace transforms of the above equations are

$$sp_0(s) - 1 = -\lambda p_0(s) + (a - ib)p_3(s) \quad (6.89)$$

$$sp_1(s) = -ap_1(s) + \lambda p_0(s) \quad (6.90)$$

$$sp_2(s) = -(a + ib)p_2(s) + ap_1(s) \quad (6.91)$$

$$sp_3(s) = -(a - ib)p_3(s) + (a + ib)p_2(s) \quad (6.92)$$

From Equations (6.90)–(6.92)

$$p_1(s) = \frac{\lambda}{s + a} p_0(s) \quad (6.93)$$

$$\begin{aligned} p_2(s) &= \frac{a}{s + a + ib} p_1(s) \\ &= \frac{a\lambda}{(s + a)(s + a + ib)} p_0(s) \end{aligned} \quad (6.94)$$

$$p_3(s) = \frac{a + ib}{s + (a - ib)} \frac{a\lambda}{(s + a)(s + a + ib)} p_0(s) \quad (6.95)$$

Substituting (6.95) into (6.89)

$$\begin{aligned} (s + \lambda)p_0(s) &= 1 + (a - ib)p_3(s) \\ &= 1 + \frac{(a - ib)(a + ib)a\lambda}{(s + a - ib)(s + a + ib)(s + a)} p_0(s) \\ \left[(s + \lambda) - \frac{(a^2 + b^2)a\lambda}{(s + a)((s + a)^2 + b^2)} \right] p_0(s) &= 1 \end{aligned}$$

Therefore

$$\begin{aligned} p_0(s) &= \frac{(s + a)((s + a)^2 + b^2)}{s\{[(s + a)^2 + b^2][s + a + \lambda] + a\lambda(s + 2a)\}} \\ p_1(s) &= \frac{\lambda((s + a)^2 + b^2)}{s\{[(s + a)^2 + b^2][s + a + \lambda] + a\lambda(s + 2a)\}} \\ p_2(s) &= \frac{a\lambda(s + a - ib)}{s\{[(s + a)^2 + b^2][s + a + \lambda] + a\lambda(s + 2a)\}} \\ p_3(s) &= \frac{a\lambda(a + ib)}{s\{[(s + a)^2 + b^2][s + a + \lambda] + a\lambda(s + 2a)\}} \\ p_{DN}(s) &= p_1(s) + p_2(s) + p_3(s) \\ &= \frac{\lambda((s + a)^2 + b^2) + a\lambda(s + a - ib) + a\lambda(a + ib)}{s\{\cdot\}} \\ &= \frac{\lambda((s + a)^2 + b^2) + a\lambda(s + 2a)}{s\{[(s + a)^2 + b^2][s + a + \lambda] + a\lambda(s + 2a)\}} \end{aligned}$$

It can be seen that $p_2(s)$ and $p_3(s)$ when inverted will yield complex probabilities but $p_{DN}(t)$ will be real.

Steady State

$$\begin{aligned} p_{DN}(t) &= \lim_{t \rightarrow \infty} sp_{DN}(s) \\ &= \frac{\lambda(a^2 + b^2) + 2a^2\lambda}{(a^2 + b^2)(a + \lambda) + 2a^2\lambda} \\ &= \frac{\lambda[a^2 + b^2 + 2a^2]}{\lambda(a^2 + b^2 + 2a^2) + a(a^2 + b^2)} \\ &= \frac{\lambda}{\lambda + \frac{a(a^2 + b^2)}{b^2 + 3a^2}} \end{aligned}$$

$$= \frac{\lambda}{\lambda + \mu}$$

$$\text{where } \mu = \frac{a(a^2 + b^2)}{b^2 + 3a^2}$$

$$\text{It can be proved that } \mu = \frac{1}{T_d}$$

where T_d = The mean down time.

References

1. D.R. Cox, The Analysis of non-Markovian Stochastic Processes by the Inclusion of Supplementary Variables, *Proc. Camb. Phil. Soc.*, 61, pp. 433–41 (1955).
2. D.R. Cox, A Use of Complex Probabilities in the Theory of Stochastic Processes, *Proc. Camb. Phil. Soc.*, 61, pp. 313–19 (1955).
3. Sheldon Ross, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco (1970).
4. C. Singh, R. Billinton, Reliability Modelling in Systems with non-Exponential Down Time Distributions, *IEEE Trans.*, PAS-92, No. 2, pp. 790–800 (March/April, 1973).
5. C. Singh, R. Billinton and S.Y. Lee, Reliability Modelling Using the Device of Stages, *PICA Proceedings* (1973).
6. C. Singh, 'Reliability Modelling and Evaluation in Electric Power Systems,' Ph.D. Thesis, University of Saskatchewan, Saskatoon, Canada (1972).
7. Electric Power Institute of Texas A & M University, *Methods of Bulk Power System Security Assessment (Probability Approach)*, Edison Electric Institute Project R P 90–6 (Nov. 1970).

CHAPTER 7

Simulation

Introduction

The previous chapters were concerned with formulating and solving mathematical models for system reliability. The solution may be either in the form of an explicit expression or the results may be obtained by numerical methods. This type of approach portrays the cause and effect relationships in the physical system and enhances the insight. Methods have been described both for Markovian and non-Markovian systems. It is obvious, however, that problems involving non-exponential distributions can become very complicated. The analytical approach is efficient and should always be employed when it is possible to develop a model which is a reasonable representation of the physical system and also when such a model is amenable to solution. Some problems are, however, too complex to be solved in this manner and simulation techniques have to be used.

In simulation, the system is divided into elements whose behaviour can be predicted either deterministically or by probability distributions. These elements are then combined to determine the system reliability. Simulation, therefore, also employs a mathematical model but it proceeds by performing sampling experiments on this mathematical model. Simulation experiments are virtually the same as ordinary statistical experiments except that they are performed on the mathematical model rather than on the actual system.

Simulation is an imprecise technique by virtue of its very statistical nature. Mathematical methods discussed in the previous chapters generally give exact results under the assumptions made. Simulation techniques, however, provide only the estimates of the exact results. Moreover they provide only a numerical value and to obtain another numerical value for a different set of parameters, the whole simulation experiment may have to be repeated. Sensitivity analysis, using a simulation approach is therefore quite expensive. It is, however, a very flexible approach and for many problems may be the only answer.

It should be noted this book deals only with digital simulation. It is called digital because most often it is executed on the digital computer but there is no inherent relationship between the two. It is a vast field and a separate book is needed to do justice to it. This book discusses the basics on which the reader can build further simulation programs.

Basic Procedure

It has been previously noted that simulation experiments are similar to ordinary statistical experiments except that they use a mathematical model of the system rather than the physical system itself. This is illustrated with the help of an example of two independent components in parallel. The system is failed when both the components are failed. The system could be constructed and operated for a long time. The history of operation and failure of the system could be recorded and the different reliability measures obtained using statistical methods. Such a method would be very expensive, especially where costly equipment is involved and would require a long time before any statement could be made about its reliability. The simulation of this system is conducted by making a mathematical model where the behaviour of the components are represented by probability distributions. Assume that component 1 is in the up state at the beginning of the experiment. Using a random number and the probability distribution of the up time of component 1, the time at which this component will fail is determined. Methods for doing this are explained later in the chapter. In a similar manner a possible duration of the repair time is generated. A history of the component generated in this manner is one possible realization of the stochastic process. The realization of component 2 is also constructed and the overlapping outage durations represent the durations of the system failure. A number of realizations of the system history can be constructed in this manner and the reliability measures obtained from these realizations using statistical methods.

In essence, simulation consists of constructing realizations of the stochastic process underlying the system and then extracting the required system performance parameters from these realizations. Most of the refinements in the theory of simulation are concerned either with developing more efficient methods of constructing realizations or extracting the information from the least possible number of realizations.

Random Number Generation

Random numbers are needed to generate random observations from the probability distributions. Tables of random numbers have been generated using mechanical or electronic devices. The basic requirement for the numbers to be random is that in a sequence each number should have equal probability of taking on any one of the possible values and it must be statistically independent of the other numbers in the sequence. While executing simulation on a digital computer, the table of random numbers can be provided externally. It is, however, more common to have the computer generate its own random numbers. There are several good methods available and only one is described here. It is a multiplicative congruential method and obtains the $(n+1)$ th random

number R_{n+1} from the n th random number R_n by using the following recurrence relation due to Lehmer

$$R_{n+1} = (aR_n) \pmod{m}$$

where a and m are positive integers, $a < m$. The above notation signifies that R_{n+1} is the remainder when (aR_n) is divided by m . The first random number R_0 is assumed and the subsequent random numbers can be generated by this recurrence relation. The sequence of random numbers so generated is periodic. Great care has to be exercised in the selection of a combination of R_0 , a and m . The sequence cycle should be larger than the number of random numbers required. One combination which is satisfactory is

$$\begin{aligned} a &= 455\,470\,314 \\ m &= 2^{31} - 1 = 2\,147\,483\,647 \\ R_0 &= \text{Any integer between 1 and } 2\,147\,483\,646. \end{aligned}$$

Now if random numbers between say 0 and 999 are required, then the computer can be instructed to take the last three digits of the random number so generated. It can be seen that the sequence of random numbers so produced is predictable and reproducible and is not therefore strictly a sequence of random numbers. For this reason these random numbers are called pseudo-random numbers. They can, however, satisfactorily play the role of random numbers in digital simulation. In fact in many applications where alternative design configurations are being evaluated, the use of the same sequence of random numbers may be desirable.

Simulation Model

A simulation model representing the system to be simulated is required. The analyst should become thoroughly familiar with the system as in the case of mathematical modelling. In complex systems, failure modes and effects analysis is quite useful in gaining an insight into the system behaviour. The system is broken into elements whose behaviour can be predicted either in a deterministic manner or in the form of probability distributions. In reliability evaluation, continuous probability distributions are most often used. When historical data is available, either the frequency distribution of these data or the probability distribution which best fits these data may be used. The latter alternative is, however, more satisfactory as it comes closer to predicting the expected future performance rather than repeating the idiosyncrasies of the recorded data.

The operating rules which define the effect of the elements on each other and on the system should be specified. These rules may be either probability distributions, tables or some set of rules. It may be preferable to draw logical

flow diagrams of the system specifying the rules and logical linkages. The tendency to be over-realistic at the expense of simplicity should be guarded against.

Timing Controls

Simulation studies deal with the passage of time. There is no connection between the simulated time which represents the passage of time in the actual world and the computational time. There are two methods of representing time in computer simulation programs:

- i. fixed time interval method
- ii. next event method

A brief description of each is given below.

Fixed Time Interval Method

This is also called the synchronous timing method. This is a two step method. The basic time interval is Δt which may be microseconds, minutes or days. The interval length Δt will be chosen depending upon the operating characteristics of the system. Starting in the initial state, time is advanced by Δt and the program then looks to see if an event has occurred. The system is then up dated by determining the resulting state of system. If no event has occurred then the system stays in the same state. These two steps may be repeated as many times as desired.

Next Event Method

This is also called the asynchronous timing method. Simulated time, in this method, is advanced by a variable amount rather than a fixed amount each time. The computer proceeds by keeping a record of the next few simulated events scheduled to occur. The most imminent event is assumed to have occurred and the simulated time is advanced to the point of occurrence of the event. The cycle is repeated as many times as desired.

In essence, in the synchronous timing method, the time is advanced by definite amounts and every time the system is updated by determining the event that occurred during this interval and in the asynchronous timing method, the next event is determined and the time is advanced to the occurrence of this event. The occurrence of an event during an interval or the time to the occurrence of an event is determined using the following sampling techniques.

Random Sampling

When all the elements operate and interact in a deterministic manner, the

event occurring during an interval or the time till the next event, is easily determined. In systems with stochastic elements, these random observations are obtained from the probability distributions using random numbers and methods of generating random observations from probability distributions are required.

Discrete Distributions

Two methods for modelling discrete distributions are described below.

1 Proportionate Allocation Technique

It consists in allocating the possible values of the random number to the various values of the random variable underlying the distribution in direct proportion to their respective probabilities. A random number is selected and the corresponding value of the random variable is the random observation. The method is quite useful in simulating discrete time Markov chains when the transition probability matrix is specified. This method is illustrated by the example of man who if he does his exercises one day, is 70% sure not to do them next day. On the other hand, if he does not do his exercises one day, is 60% sure not to do them the next day. Denoting the doing and not doing of the exercises by 0 and 1 respectively, the transition probability matrix is

Initial state	Final state	
	0	1
0	0.3	0.7
1	0.4	0.6

The realizations are constructed using a table of random digits. If the man is in state 0, select a single random digit and the next state is determined as follows

digit	event
0 - 2	stay in 0
3 - 9	transit to 1

Similarly if the man is in state 1

digit	event
0 - 3	transit to 0
4 - 9	stay in 1

The construction of a realization for ten days is shown in Table 1. It is assumed that the man does his exercises on the first day. It should be noted that this is only one possible realization of the stochastic process. The stochastic nature is evident from the sequence of states occupied. For deriving probabilities,

a number of such realizations have to be constructed. The methods for deriving measures from these realizations is described later in the chapter.

Table 7.1 Constructing a Realization

Day	Random number	State
1		0
2	4	1
3	3	0
4	4	1
5	1	0
6	2	0
7	2	0
8	2	0
9	2	0
10	0	0

2 Inverse of the Probability Distribution Method

This method is practically the same as method 1 but is a little more involved and proceeds in the following steps:

1. Construct the distribution function of the random variable X , i.e. $F(x) = P(X \leq x)$. The distribution function has the property that is monotonically increasing. The probability mass function of an arbitrary random variable X and its distribution function are shown in Fig. 7.1.

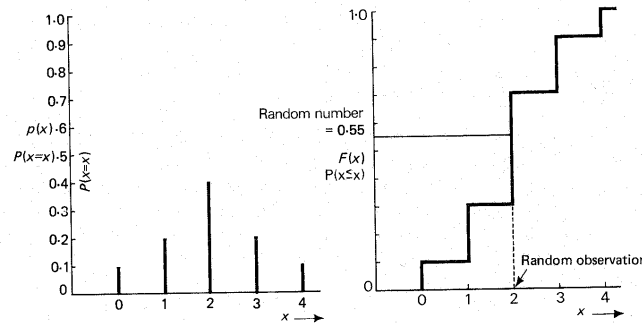


Fig. 7.1 Probability mass function and probability distribution function

2. Generate a random decimal number between 0 and 1. This is achieved by obtaining a random integer with the desired digits and then placing a decimal point before it.
3. Set $F(x)$ equal to the random number and select the value of x corresponding to $F(x)$. This value of x is the desired random observation from the probability distribution.

It should be noticed that $F(x_i) - F(x_{i-1})$ is equal to $P(X = x_i)$, and if the random number falls in the interval $(F(x_i), F(x_{i+1}))$, the value of $X = x_i$ will be selected. The procedure therefore essentially allocates the random numbers to the random variables in the proportion of their probabilities of occurrence. The basic procedure is therefore the same as that of method 1.

Continuous Distributions

The above procedure can also be used for continuous distributions. Continuous distributions are approximated by discrete distributions whose irregularly spaced points have equal probabilities. The accuracy can be increased by increasing the number of intervals into which (0,1) is divided. This requires additional data in the form of tables. Although the method is quite general, its disadvantages are the great amount of work required to develop tables and possible computer storage problems. The following analytic inversion approach is simpler.

Let z be a random number in the range 0 to 1 with either a uniform probability density function or a triangular distribution function, i.e.

$$f(z) = \begin{cases} 0 & Z < 0 \\ 1 & 0 \leq Z \leq 1 \\ 0 & Z > 1 \end{cases}$$

Similarly

$$F(z) = \begin{cases} 0 & Z < 0 \\ z & 0 \leq Z \leq 1 \\ 1 & Z \geq 1 \end{cases}$$

Let $F(x)$ be the distribution function from which the random observations are to be generated. Let

$$z = F(x)$$

Solving the equation for x gives a random observation of X . That the observations so generated do have $F(x)$ as the probability distribution can be shown as follows.

Let ϕ be the inverse of F , then

$$x = \phi(z)$$

Now x is the random observation generated. To determine its probability distribution

$$P(x \leq X) = P(F(x) \leq F(X)) = P(z \leq F(X)) = F(X)$$

Therefore the distribution function of x is $F(x)$ as required. In the case of several important distributions, special techniques have been developed for efficient random sampling. A few cases are described in the following section. The reader can refer to Reference 5 for a more detailed treatment.

Exponential Distribution

The exponential distribution has the following probability distribution

$$P(X \leq x) = 1 - e^{-\rho x}$$

where $1/\rho$ is the mean of the random variable X . Setting this function equal to a random decimal number between 0 and 1

$$z = 1 - e^{-\rho x}$$

Since the complement of such a random number is also a random number the above equation can be written as

$$z = e^{-\rho x}$$

Taking the natural logarithm of both sides and simplifying

$$x = \frac{\ln(z)}{-\rho}$$

which is the desired random observation from the exponential distribution having $1/\rho$ as the mean.

Erlang Distribution

The above procedure can be readily extended to generate random observations from an Erlang distribution. It has been shown in Chapter 6 that the sum of ' a ' independent exponentially distributed random variables each with mean $\frac{1}{\rho}$ has the Erlang distribution with mean $\frac{a}{\rho}$ and ' a ' as the shape parameter.

Therefore if we have a sequence of ' a ' random decimal numbers in the interval (0,1), denoted by $z_1, z_2, \dots, z_a$, the random observation from Erlang distribution is

$$x = \sum_{i=1}^a \frac{\ln(z_i)}{-\rho}$$

$$= -\frac{1}{\rho} \left(\ln \prod_{i=1}^a z_i \right)$$

Normal Distribution

The following describes a technique developed by Box and Muller. The method proceeds by generating pairs of normal deviates. The joint distribution of two independent standardized normal deviates is given by

$$f_{XY}(x, y) = \frac{1}{2\pi} \exp \left\{ -\frac{1}{2}(x^2 + y^2) \right\} \quad (7.1)$$

Consider the polar transformation

$$x = u \cos v$$

$$y = u \sin v$$

The inverse transformation is

$$u = (x^2 + y^2)^{\frac{1}{2}}$$

and

$$v = \arctan \frac{y}{x}$$

The above functions can be written in a general form

$$U = U(X, Y) \quad V = V(X, Y)$$

and

$$X = X(U, V) \quad Y = Y(U, V)$$

When X is near x and Y is near y , U and V must be near u and v . Therefore

$$P(x < X \leq x + dx, y < Y \leq y + dy) = f_{XY}(x, y) dx dy$$

$$= P(u < U \leq u + du, v < V \leq v + dv)$$

$$= f_{UV}(u, v) du dv \quad (7.2)$$

Therefore

$$f_{UV}(u, v) = f_{XY}(x, y) \left| \frac{dx}{du} \frac{dy}{dv} \right|$$

Absolute values are used so that the expression is applicable for both non-increasing and non-decreasing functions.

$$dx dy = u du dv$$

Therefore

$$\begin{aligned} f_{UV}(u, v) du dv &= \frac{1}{2\pi} e^{-u^2} u du dv \\ &= e^{-u^2} d\left(\frac{1}{2}u^2\right) \frac{1}{2\pi} dv \end{aligned} \quad (7.4)$$

This expression can be interpreted in this manner: u^2 is exponentially distributed and v has a uniform distribution in the interval $(0, 2\pi)$. Now if there are two random decimal numbers z_1, z_2 in $(0, 1)$

$$u^2 = -\ln z_1$$

i.e.

$$u = \sqrt{\ln 1/z_1}$$

and

$$v = 2\pi z_2$$

Hence

$$X = \sqrt{\ln \frac{1}{z_1}} \cos 2\pi z_2$$

and

$$Y = \sqrt{\ln \frac{1}{z_1}} \sin 2\pi z_2$$

are exact independent normal deviates.

This book covers only a few cases of analytic inversion. Many other methods exist for analytic inversion of particular probability distributions including inversion by graphical or tabular means.

Estimating Reliability Measures

Reliability measures can be calculated from the realizations using statistical methods. It is possible to construct probability distributions for the various residence times but usually the mean values are the main parameters.

Time Specific Probability of X^+

If N observations of the state of system are made at time t , and n_+ of the times the system is found in X^+ , the estimate of the probability of X^+ is found by

$$\hat{p}_+(t) = \frac{n_+}{N} \quad (7.5)$$

It is well known from the frequency concept of probability that

$$p_+(t) = \hat{p}_+(t) = \frac{n_+}{N} \quad N \rightarrow \infty$$

Interval Frequency

If in N realizations of the system state, the system visited X^+ n_c times, the estimate of $F(0, t)$ is

$$\hat{F}_+(0, t) = \frac{n_c}{N} \quad (7.6)$$

Fractional Duration

The fractional duration in X^+ is estimated by

$$\hat{D}_+(0, t) = \frac{\sum_{i=1}^N u_i}{tN} \quad (7.7)$$

where u_i is the time spent in X^+ in the i th realization.

Steady State Probability of being in X^+

This can be calculated either from a number of realizations, allowing sufficient time for letting each realization to reach equilibrium state or it can be calculated as the fraction of time spent in X^+ in a very long realization, i.e.

$$\hat{P}_+ = \frac{t_+}{T}$$

where t_+ is the time spent in X^+ in interval $(0, T)$ when T is large. This latter approach is less time consuming as in the former approach considerable time is spent in reaching the equilibrium condition.

Mean Cycle Time

After the simulated system has reached an equilibrium, the mean cycle time can be estimated as follows

$$\hat{T}_+ = \frac{\sum_{i=1}^n t_i}{n}$$

where t_i is the time interval between the $(i-1)$ th and the i th encounter of X^+ .

Equilibrium Conditions and Sample Size

The most important reliability measures in repairable systems are obtained from the steady state or equilibrium conditions. In this case the system state probabilities are independent of the initial conditions and the time elapsed since the start of system. The system reaches a steady state condition when the system state probability distribution reaches a limiting equilibrium distribution. It should be remembered that steady state condition can only be approached but never exactly attained.

In determining the mean cycle time, the data obtained during the initial period of simulated system operation should be excluded. It is difficult to know how long the system should be operated before taking observations. It could, however, be roughly estimated by having a few trial runs and estimating the probability distributions at various points in time. It should be noted that even when steady state measures are the basis of evaluating alternative designs, the same initial conditions should preferably be used. Simulation or Monte Carlo techniques are used to obtain a numerical estimate of the inherent system reliability parameters. As the sample size increases the estimated value approaches the estimand. The question which now arises is how big should the sample size be? It is not adequate to simulate the system for an arbitrary long time and then simply assume that the results are sufficiently precise.

Reliability measures obtained from each simulated sample run are generally different and one object is to determine the mean value of the measure. A simple case is when these observations of the measure are statistically independent and have a common normal distribution. This case is considered and then extended to more general situations. The confidence interval for the mean m of the normal distribution having variance σ^2 can be obtained as follows. Let $\bar{X}$ be the sample mean obtained from a random sample of size n . $\bar{X}$ is then also a random variable and it can be shown that

$$Z = \frac{(\bar{X} - m)\sqrt{n}}{\sigma}$$

has a t distribution with $(n - 1)$ degrees of freedom. The value of σ is

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

where X_i is the i th observation. Now

$$P\left\{-t_{\alpha/2}^{n-1} \leq \frac{(\bar{X} - m)\sqrt{n}}{\sigma} \leq t_{\alpha/2}^{n-1}\right\} = 1 - \alpha$$

where $t_{\alpha/2}^{n-1}$ is the 100 $\alpha/2$ percent point of the distribution with $(n - 1)$ degrees of freedom and can be found from Tables of the t distribution. The above

expression can be rearranged as

$$P\left\{\bar{X} - t_{\alpha/2}^{n-1} \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + t_{\alpha/2}^{n-1} \frac{\sigma}{\sqrt{n}}\right\} = 1 - \alpha$$

Therefore with confidence $1 - \alpha$, the upper and lower bounds of m are

$$\text{upper bound} = \bar{X} + t_{\alpha/2}^{n-1} \frac{\sigma}{\sqrt{n}}$$

$$\text{lower bound} = \bar{X} - t_{\alpha/2}^{n-1} \frac{\sigma}{\sqrt{n}}$$

The required sample size can be predicted by obtaining an estimate of the standard deviation of the observations either from pilot runs or from initial observations. As can be seen from above, the interval between the upper and lower bounds can be made as narrow as desired by making the sample size sufficiently large.

Two assumptions made above are that the observations of the measure are statistically independent and have a common normal distribution. This, however, is not true in general and the following methods may be employed to realize these assumptions in practice.

The first method is to have a number of independent simulated runs. The mean measure obtained from each run can be used as an independent observation. These can be assumed to be normally distributed in accordance with the Central Limit Theorem. The confidence interval therefore can be found by the procedure described.

When steady state measures are being estimated, the above procedure can be quite wasteful since in each simulation run, the initial period is unproductive. The alternative is to use a single simulated run and to divide the steady state period into equal long intervals. The value derived from each interval can be used as an observation. It should be noted that these observations are not completely independent but by making the intervals sufficiently long, the correlation can be decreased.

Variance Reducing Techniques

It has been pointed out that the precision of sample estimates can be increased by making the sample size large enough. Increasing the precision is equivalent to decreasing the variance of the sample estimates. The simple method of repeated runs (or making a single run very large and dividing it into equal intervals), treating the measures obtained from each sample as independent sample values until the variance has been reduced to the desired level is usually quite time-consuming. Special techniques for reducing variance have

been devised. These techniques, extract as much and as precise information as possible from the amount of simulation that can be economically executed. These are three generally used techniques:

1. stratification
2. control variates
3. antithetic variables

The reader is referred to References 4–5 for discussion of these methods.

References

1. D.R. Cox and H.D. Miller, *The Theory of Stochastic Processes*, Methuen, London (1965).
2. C.D. Flagle, W.H. Huggins and R.H. Roy, *Operations Research and Systems Engineering*, Johns Hopkins Press, Baltimore (1960).
3. F.S. Hillier and G.J. Lieberman, *Introduction To Operations Research*, Holden-Day (1970).
4. J.M. Hammersley and D.C. Handscomb, *Monte Carlo Methods*, Methuen, London (1964).
5. K.D. Tocher, *The Art of Simulation*, English Universities Press, London (1963).
6. C.F. DeSieno and L.L. Stine, A Probability Method for Determining the Reliability of Electric Power Systems, *Trans. AIEE* (February 1964).

CHAPTER 8

Conclusions

We began this book by describing the place of modelling and evaluation in the total setting of system reliability planning. This is an important phase of a reliability program as the reliability model integrates the information available on the various components to provide overall system indices. Several reliability measures have been defined and techniques described for their calculation. The choice of a particular index depends on the penalty factors associated with system failures. If the penalty depends on the total duration of failure, availability is a relevant measure. If, however, the number of failures is more important, frequency is a more appropriate index. Ideally a reliability measure should be sensitive to all the factors which affect reliability. In practice, however, some measures are more sensitive to a certain parameter than the others. A single measure is, therefore, not likely to give a complete picture of system reliability and the use of more than one reliability index is preferable.

The central concept running throughout the book is the *frequency balancing approach*. This method gives the same linear, differential or integro-differential equations as would be obtained by conventional means. The chief advantage of this approach is the simplicity with which frequency, cycle time and mean duration indices can be calculated. Another important feature of this technique is the equivalent transition rate. This concept has been used to derive conditions of mergeability which set the mathematical limits for model reduction. Correct appreciation of the conditions of mergeability is important for model reduction. Many gross mistakes have been committed in the literature because of a lack of such an understanding.

Reliability evaluation of large and complex systems can sometimes be quite difficult. The difficulties are more computational than conceptual. When the system is composed of independent elements, model reduction using the concept of equivalent transfer rates proves quite useful. When dependent failures are encountered, model reduction becomes limited. Two useful techniques under these circumstances are truncation and sequential truncation.

In reliability models, assumptions are made about the probability distributions of times to failure and times to restore. It goes without saying that model formulation approaches reality only to the extent that the assumed distributions approach the actual ones. The bulk of this book is devoted to Markovian models. These models assume exponential distributions for the mean up and down times. The techniques of dealing with non-Markovian models have also been described

in detail. Analytical methods involving integro-differential equations are useful but the solution can be quite intractable for fairly complex systems. The device of stages transforms a non-Markovian model into a Markovian model and a solution is therefore possible. The only disadvantage of this method is the multiplication of the number of states.

Mathematical modelling provides a compact and efficient means of obtaining system reliability measures. The mathematical model abstracts the essence of the physical system and provides a deeper insight into the cause and effect relationships within the system. All reliability problems, however, cannot be solved analytically. Simulation is a more flexible approach and can take into account many factors which cannot be handled by mathematical models. The approach, however, is time-consuming and does not provide the same insight into the system. Wherever possible, an analytical approach is definitely superior to using a simulation technique.

The emphasis throughout the book is not on the development of particular reliability models but on deepening the insight into the theory, philosophy and the techniques of system reliability modelling. The purpose is to enhance the reader's capability for reliability modelling and evaluation.

We conclude this book by repeating the warning given in the introduction. After making a number of experiments with the reliability model, it is possible to develop a confidence in the results which may not prove justified by a closer examination of the data. The reliability model transforms data into reliability measures and, given a correct model, the validity of these measures depends on the validity of the data used.

CHAPTER 9

Appendices

Appendix I

Solution of simultaneous linear equations

Reliability evaluation often calls for the solution of a set of simultaneous linear equations of the form

$$AX = B$$

where A is a nonsingular coefficient matrix and X and B are column vectors. The results could be obtained using Cramer's rule which proceeds by evaluating determinants and expanding by minors. The system of n equations in n unknowns takes on the order of $n!$ multiplications. If $n = 25$ and each multiplication takes 10^{-6} seconds, the computation time required would be several million years. Several numerical methods have been devised to solve linear equations. This book covers only the basic principles of Gauss-Jordan method of elimination. Readers interested in further details should refer to one of the many excellent books available on numerical methods.

The basic procedure of this method is quite simple. The first variable from all but one of the equations is eliminated by adding an appropriate multiple of this equation to each of the others. The second variable is then eliminated from another equation in the same manner. The procedure continues until each of the equations has only one variable left. The result can then be read directly. This can be illustrated by solving the following set of linear equations.

$$x_1 + x_2 + 3x_3 = 4 \quad (1)$$

$$x_1 - x_2 + 4x_3 = 5 \quad (2)$$

$$2x_1 - x_2 + 3x_3 = 6 \quad (3)$$

Step 1 Remove x_1 from (2) and (3) by multiplying Equation (1) by (-1) and (-2) respectively and adding

$$x_1 + x_2 + 3x_3 = 4 \quad (4)$$

$$-2x_2 + x_3 = 1 \quad (5)$$

$$-3x_2 - 3x_3 = -2 \quad (6)$$

Step 2 Remove x_2 from (4) and (6) by multiplying (5) by $(\frac{1}{2})$ and $(-\frac{3}{2})$ respectively and adding to (4) and (6) respectively

$$x_1 + \frac{7}{2}x_3 = \frac{9}{2} \quad (7)$$

$$-2x_2 + x_3 = 1 \quad (8)$$

$$-\frac{9}{2}x_3 = -\frac{7}{2} \quad (9)$$

Step 3 Remove x_3 from (7) and (8) by multiplying (9) by $(\frac{2}{9})$ and $(\frac{2}{9})$ and adding to (7) and (8) respectively

$$x_1 = \frac{16}{9}$$

$$-2x_2 = \frac{2}{9}$$

$$-\frac{9}{2}x_3 = -\frac{7}{2}$$

Step 4 The results can now be read

$$x_1 = \frac{16}{9}$$

$$x_2 = -\frac{1}{9}$$

$$x_3 = \frac{7}{9}$$

In the computer, the operations are done in matrix form. The initial step is to form an augmented array

$$[A \mid B]$$

This in our present example is of the form

$$\begin{bmatrix} 1 & 1 & 3 & 4 \\ 1 & -1 & 4 & 5 \\ 2 & -1 & 3 & 6 \end{bmatrix}$$

Row 1 is called the pivot row and the first element of this row is called the pivot element. The procedure consists in setting the elements above and below the pivot elements (diagonal elements) to zero.

First step. R_1 (row 1) is the pivot row, a_{11} is the pivot element. Divide the pivot row by the pivot element. Set elements below the pivot element to zero by

$$R'_2 = R_2 - a_{21}R_1$$

$$R'_3 = R_3 - a_{31}R_1$$

$$\begin{bmatrix} 1 & 1 & 3 & 4 \\ 0 & -2 & 1 & 1 \\ 0 & -3 & -3 & -2 \end{bmatrix}$$

The prime indicates the resulting row.

Second Step. R_2 is the pivot row and a_{22} is the pivot element. Normalize the pivot row by dividing by the pivot element.

$$\begin{bmatrix} 1 & 1 & 3 & 4 \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -3 & -3 & -2 \end{bmatrix}$$

Set the elements above and below the pivot element by

$$R'1 = R1 - a_{12}R2$$

$$R'3 = R3 - a_{32}R2$$

$$\begin{bmatrix} 1 & 0 & \frac{7}{2} & \frac{9}{2} \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{9}{2} & -\frac{7}{2} \end{bmatrix}$$

Third Step. R_3 is the pivot row and a_{33} is the pivot element. Normalize the pivot row by dividing by the pivot element.

$$\begin{bmatrix} 1 & 0 & \frac{7}{2} & \frac{9}{2} \\ 0 & 1 & -\frac{1}{2} & -\frac{7}{2} \\ 0 & 0 & 1 & \frac{7}{9} \end{bmatrix}$$

Set the elements above the pivot row by

$$R'1 = R1 - a_{13}R3$$

$$R'2 = R2 - a_{23}R3$$

$$\begin{bmatrix} 1 & 0 & 0 & \frac{16}{9} \\ 0 & 1 & 0 & -\frac{1}{9} \\ 0 & 0 & 1 & \frac{7}{9} \end{bmatrix}$$

The last column is the solution vector. It should be noted that special pivoting techniques are available for avoiding rounding off errors. For details of further refinements, books on numerical analysis should be consulted.

Appendix II

Shape of the hazard rate function of two series stage combinations in parallel

The expression of the probability density function of this stage combination as given in Chapter 6 is

$$f(x) = \omega_1 \rho_1 \frac{(\rho_1 x)^{a_1-1}}{(a_1-1)!} e^{-\rho_1 x} + \omega_2 \rho_2 \frac{(\rho_2 x)^{a_2-1}}{(a_2-1)!} e^{-\rho_2 x} \quad (1)$$

The survivor function is

$$S(x) = \omega_1 \sum_{n=1}^{a_1} \frac{(\rho_1 x)^{n-1}}{(n-1)!} e^{-\rho_1 x} + \omega_2 \sum_{n=1}^{a_2} \frac{(\rho_2 x)^{n-1}}{(n-1)!} e^{-\rho_2 x} \quad (2)$$

and the hazard rate function is

$$\phi(x) = \frac{f(x)}{S(x)}$$

(i) At the origin

$$\begin{aligned} \phi(0) &= f(0) \\ \text{since } S(0) &= 1 \end{aligned}$$

The following conclusions can be drawn regarding the magnitude of the hazard rate at the origin

$$\phi(0) = \begin{cases} 0 & \text{if } a_1 > 1, \quad a_2 > 1 \\ \omega_1 \rho_1 & \text{if } a_1 = 1, \quad a_2 > 1 \\ \omega_2 \rho_2 & \text{if } a_1 > 1, \quad a_2 = 1 \\ \omega_1 \rho_1 + \omega_2 \rho_2 & \text{if } a_1 = a_2 = 1 \end{cases}$$

(ii) The derivative of $\phi(x)$ at the origin

$$\begin{aligned} \phi'(0) &= \frac{f'(x)S(x) + \{f(x)\}^2}{\{S(x)\}^2} \Big|_{x=0} \\ &= f'(0) + \{f(0)\}^2 \end{aligned}$$

Now by the initial value theorem

$$f'(0) = \lim_{s \rightarrow \infty} s[s\bar{f}(s) - f(0)]$$

The expression for $\bar{f}(s)$ is

$$\bar{f}(s) = \omega_1 \left(\frac{\rho_1}{\rho_1 + s} \right)^{a_1} + \omega_2 \left(\frac{\rho_2}{\rho_2 + s} \right)^{a_2}$$

The following conclusions can be drawn

(a) $a_1 = a_2 = 1$

Then

$$\begin{aligned} s\bar{f}(s) - f(0) &= \frac{s\omega_1 \rho_1}{s + \rho_1} + \frac{s\omega_2 \rho_2}{s + \rho_2} - (\omega_1 \rho_1 + \omega_2 \rho_2) \\ &= -\frac{\omega_1 \rho_1^2}{\rho_1 + s} - \frac{\omega_2 \rho_2^2}{\rho_2 + s} \end{aligned}$$

Therefore

$$\begin{aligned} f'(0) &= \lim_{s \rightarrow \infty} s \left[-\frac{\omega_1 \rho_1^2}{\rho_1 + s} - \frac{\omega_2 \rho_2^2}{\rho_2 + s} \right] \\ &= -\omega_1 \rho_1^2 - \omega_2 \rho_2^2 \end{aligned}$$

and

$$\begin{aligned} \phi'(0) &= -\omega_1 \rho_1^2 - \omega_2 \rho_2^2 + (\omega_1 \rho_1 + \omega_2 \rho_2)^2 \\ &= -\omega_1 \omega_2 (\rho_1 - \rho_2)^2 \end{aligned}$$

This expression is always negative and therefore for this condition, the hazard rate is always initially decreasing. See curve 1 of Fig. 6.9.

(b) $a_1 = 1$ and $a_2 = 2$

$$s\bar{f}(s) - f(0) = -\frac{\omega_1 \rho_1^2}{s + \rho_1} + \omega_2 s \left(\frac{\rho_2}{s + \rho_2} \right)^2$$

Therefore

$$f'(0) = \lim_{s \rightarrow \infty} s \{s\bar{f}(s) - f(0)\} = -\omega_1 \rho_1^2 + \omega_2 \rho_2^2$$

and

$$\begin{aligned} \phi'(0) &= -\omega_1 \rho_1^2 + \omega_2 \rho_2^2 + (\omega_1 \rho_1)^2 \\ &= \omega_2 (\rho_2^2 - \omega_1 \rho_1^2) \end{aligned}$$

The sign will depend upon that of the quantity inside the parenthesis. This is negative for curve 2 in Fig. 6.9 and therefore the hazard rate is initially decreasing.

(c) $a_1 = 2$ and $a_2 = 1$

$$\phi'(0) = \omega_1 (\rho_1^2 - \omega_2 \rho_2^2)$$

$$(d) \quad a_1 = 2 \quad \text{and} \quad a_2 = 2$$

$$\phi'(0) = f'(0) \quad \text{since} \quad f(0) = 0$$

Therefore

$$\begin{aligned} \phi'(0) &= \lim_{s \rightarrow \infty} s^2 \bar{f}(s) \\ &= \omega_1 \rho_1^2 + \omega_2 \rho_2^2 \end{aligned}$$

The hazard rate is therefore initially increasing in this case. See curve 4 in Fig. 6.9

$$(e) \quad a_1 > 2, \quad a_2 > 2$$

$$\begin{aligned} \phi'(0) &= \lim_{s \rightarrow \infty} s^2 \bar{f}(s) \\ &= 0 \end{aligned}$$

The hazard rate is, therefore, initially constant as can be verified from curve 5 of Fig. 6.9

$$(f) \quad a_1 = 1 \quad \text{and} \quad a_2 > 2$$

$$\begin{aligned} s \bar{f}(s) - f(0) &= -\frac{\omega_1 \rho_1^2}{\rho_1 + s} + \omega_2 s \left(\frac{\rho_2}{s + \rho_2} \right)^{a_2} \\ \lim_{s \rightarrow \infty} s \{s \bar{f}(s) - f(0)\} &= -\omega_1 \rho_1^2 \end{aligned}$$

Therefore

$$\begin{aligned} \phi'(0) &= -\omega_1 \rho_1^2 + (\omega_1 \rho_1)^2 \\ &= -\omega_1 \omega_2 \rho_1^2 \end{aligned}$$

$$(g) \quad a_1 > 2 \quad \text{and} \quad a_2 = 1$$

$$\phi'(0) = -\omega_1 \omega_2 \rho_2^2$$

In cases (f) and (g) the hazard rate is, therefore, initially decreasing as can be seen from curve 3 of Fig. 6.9.

$$(III) \quad \phi(x) \quad \text{as} \quad x \rightarrow \infty$$

$$(a) \quad \rho_1 > \rho_2$$

$$\lim_{x \rightarrow \infty} \phi(x) = \rho_2$$

$$(b) \quad \rho_2 > \rho_1$$

$$\lim_{x \rightarrow \infty} \phi(x) = \rho_1$$

$$(c) \quad \rho_1 = \rho_2$$

$$\lim_{x \rightarrow \infty} \phi(x) = \rho_1 = \rho_2$$

The limiting value of the hazard rate as x becomes large is therefore always the smaller of ρ_1 or ρ_2 .

The three quantities

$$i. \phi(x) \quad \text{as} \quad x \rightarrow 0$$

$$ii. \phi'(x) \quad \text{as} \quad x \rightarrow 0$$

and

$$iii. \phi(x) \quad \text{as} \quad x \rightarrow \infty$$

are enough to get an approximate idea of the shape of the hazard rate. A knowledge of the behaviour of these quantities is helpful in making finer adjustments in the shape of the hazard rate.

Appendix III

Hazard rate shape of series stages in series with a distinctive stage

$$(i) \quad \phi(x) \quad \text{at} \quad x = 0$$

It can be seen by examining the ratio of Equations (6.69) and (6.70) that

$$\phi(0) = 0$$

$$(ii) \quad \phi'(x) \quad \text{at} \quad x = 0$$

As in Appendix II

$$\begin{aligned} \phi'(0) &= \lim_{s \rightarrow \infty} s [s \bar{f}(s) - f(0)] + \{f(0)\}^2 \\ &= \lim_{s \rightarrow \infty} s^2 \bar{f}(s) \quad \text{since} \quad f(0) = 0 \end{aligned}$$

Now the Laplace transform of Equation (6.69) is

$$\bar{f}(s) = \left(\frac{\rho}{\rho + s} \right)^a \frac{1}{\rho_1 + s}$$

Therefore

$$\phi'(0) = \begin{cases} \rho \rho_1 & \text{if } a = 1 \\ 0 & \text{if } a > 1 \end{cases}$$

(iii) $\phi(x)$ as $x \rightarrow \infty$

It can be proved by examining the ratio of Equation (6.69) and (6.70) that

$$\lim_{x \rightarrow \infty} \phi(x) = \begin{cases} \rho_1 & \text{if } \rho_1 < \rho \\ \rho & \text{if } \rho < \rho_1 \end{cases}$$

The final value is therefore the lesser of the two.

Appendix IV

Series stages in series with two parallel stages

Derivation of expressions

Designating the state whose duration is represented by this combination as 0 and the state of not being in it as A the state transition diagram is shown in Fig. IV.1

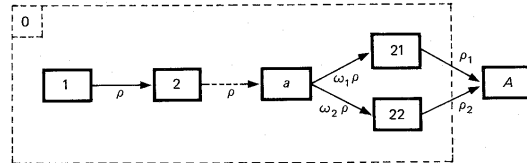


Fig. IV.1 State transition diagram to derive expressions for the stage combination

Assuming $p_1(0) = 1.0$, the time spent in state 0 is identical with the time since the origin and as explained in Chapter 6.

$$f_0(x) = \rho_1 p_{21}(x) + \rho_2 p_{22}(x)$$

$$S_0(x) = \sum_{i=1}^a p_i(x) + p_{21}(x) + p_{22}(x)$$

and

$$\phi(x) = \frac{f_0(x)}{S_0(x)}$$

The differential equations for this system are

$$p'_1(t) = -\rho p_1(t)$$

$$p'_2(t) = \rho(p_1(t) - p_2(t))$$

$$\vdots$$

$$p'_n(t) = \rho(p_{n-1}(t) - p_n(t))$$

$$\vdots$$

$$p'_a(t) = \rho(p_{a-1}(t) - p_a(t))$$

$$p'_{21}(t) = \omega_1 \rho p_a(t) - p_{21}(t) \rho_1$$

$$p'_{22}(t) = \omega_2 \rho p_a(t) - p_{22}(t) \rho_2$$

Taking the Laplace transform of these differential equations

$$p_1(s) = \frac{1}{\rho + s} \quad (1)$$

$$p_2(s) = \frac{\rho}{(\rho + s)^2}$$

Similarly

$$p_n(s) = \frac{\rho^{n-1}}{(\rho + s)^n} \quad (2)$$

$$p_a(s) = \frac{\rho^{a-1}}{(\rho + s)^a} \quad (3)$$

$$p_{21}(s) = \omega_1 \left(\frac{\rho}{\rho + s} \right)^a \frac{1}{s + \rho_1} \quad (4)$$

and

$$p_{22}(s) = \omega_2 \left(\frac{\rho}{\rho + s} \right)^a \frac{1}{s + \rho_2} \quad (5)$$

The inverse of Expression (2) is

$$p_n(x) = \frac{(\rho x)^{n-1}}{(n-1)!} e^{-\rho x} \quad (6)$$

For taking the inverse of Expression (4), it can be expanded into partial fractions

$$p_{21}(s) = \omega_1 \left(\frac{\rho}{\rho + s} \right)^a \frac{1}{s + \rho_1} = \frac{N_1}{\rho + s} + \frac{N_2}{(\rho + s)^2} + \dots + \frac{N_a}{(\rho + s)^a} + \frac{M}{\rho_1 + s}$$

The numerators can be determined as

$$M = (\rho_1 + s)p_{21}(s)|_{s=-\rho_1} = \omega_1 \left(\frac{\rho}{\rho - \rho_1} \right)^a$$

$$N_a = (\rho + s)p_{21}(s)|_{s=-\rho} = \omega_1 \rho^a \frac{1}{\rho - \rho_1}$$

and

$$N_{a-m} = \frac{1}{m!} \frac{d^m}{ds^m} (\rho + s)p_{21}(s)|_{s=-\rho} = (-1)^m \omega_1 \rho^a \frac{1}{(\rho_1 - \rho)^{m+1}}$$

Let

$$a - m = i$$

Then

$$N_i = (-1)^{a-i} \omega_1 \rho^a \frac{1}{(\rho_1 - \rho)^{a-i+1}}$$

Substituting these values into the expansion of $p_{21}(s)$ and inverting

$$\begin{aligned} p_{21}(x) &= \omega_1 \left(\frac{\rho}{\rho - \rho_1} \right)^a e^{-\rho_1 x} - \omega_1 \rho^a \frac{1}{(\rho - \rho_1)^a} e^{-\rho x} - \dots \\ &\quad - \omega_1 \rho^a \frac{1}{(\rho - \rho_1)^{a-i+1}} \frac{x^{i-1}}{(i-1)!} e^{-\rho x} \\ &\quad - \dots - \omega_1 \rho^a \frac{1}{\rho - \rho_1} \frac{x^{a-1}}{(a-1)!} e^{-\rho x} \end{aligned}$$

A similar expression can be now easily derived from $p_{22}(x)$ and finally

$$\begin{aligned} S_0(x) &= p_0(x) = \sum_{n=1}^a \frac{(\rho x)^{n-1}}{(n-1)!} e^{-\rho x} + \omega_1 \left(\frac{\rho}{\rho - \rho_1} \right)^a \\ &\quad \times \left[e^{-\rho_1 x} - e^{-\rho x} \sum_{n=1}^a \frac{\{(\rho - \rho_1)x\}^{n-1}}{(n-1)!} \right] \\ &\quad + \omega_2 \left(\frac{\rho}{\rho - \rho_2} \right)^a \left[e^{-\rho_2 x} - e^{-\rho x} \sum_{n=1}^a \frac{\{(\rho - \rho_2)x\}^{n-1}}{(n-1)!} \right] \end{aligned} \quad (7)$$

and

$$\begin{aligned} f_0(x) &= p_{21}(x)\rho_1 + p_{22}(x)\rho_2 \\ &= \omega_1 \rho_1 \left(\frac{\rho}{\rho - \rho_1} \right)^a \left[e^{-\rho_1 x} - e^{-\rho x} \sum_{n=1}^a \frac{\{(\rho - \rho_1)x\}^{n-1}}{(n-1)!} \right] \\ &\quad + \omega_2 \rho_2 \left(\frac{\rho}{\rho - \rho_2} \right)^a \left[e^{-\rho_2 x} - e^{-\rho x} \sum_{n=1}^a \frac{\{(\rho - \rho_2)x\}^{n-1}}{(n-1)!} \right] \end{aligned} \quad (8)$$

The Expression (8) could alternatively be derived using the fact that the Laplace transform of the sum of independent random variables is the product of their Laplace transforms. The probability density function of the random variable X representing the state 0 is the sum of the random variables X_1, X_2

denoting the states 1 to a and 21 and 22. The probability density function of stages 1 to a is

$$f_1(x) = \frac{\rho(\rho x)^{a-1}}{(a-1)!} e^{-\rho x}$$

and

$$\bar{f}_1(s) = \left(\frac{\rho}{\rho + s} \right)^a$$

The probability density function of the parallel stages is

$$f_2(x) = \omega_1 \rho_1 e^{-\rho_1 x} + \omega_2 \rho_2 e^{-\rho_2 x}$$

and the corresponding Laplace is

$$\bar{f}_2(s) = \omega_1 \frac{\rho_1}{\rho_1 + s} + \omega_2 \frac{\rho_2}{\rho_2 + s}$$

Now

$$\bar{f}_0(s) = \bar{f}_1(s) \cdot \bar{f}_2(s)$$

Substituting the values and inverting, $f_0(x)$ can be obtained. This is left as an exercise for the reader.

The mean duration = The sum of the means

$$= \frac{a}{\rho} + \frac{\omega_1}{\rho_1} + \frac{\omega_2}{\rho_2}$$

Variance

The two random variables X_1 and X_2 are independent and therefore the variance of X is the sum of the variances of X_1 and X_2 .

$$\begin{aligned} \text{Variance} &= \frac{a}{\rho^2} + \frac{2\omega_1}{\rho_1^2} + \frac{2\omega_2}{\rho_2^2} - \left(\frac{\omega_1}{\rho_1} + \frac{\omega_2}{\rho_2} \right)^2 \\ &= \frac{a}{\rho^2} + 2 \left(\frac{\omega_1}{\rho_1^2} + \frac{\omega_2}{\rho_2^2} \right) - \left(\frac{\omega_1}{\rho_1} + \frac{\omega_2}{\rho_2} \right)^2 \end{aligned}$$

Shape Of The Hazard Rate

$$(i) \quad \phi(0) = \frac{f(0)}{S(0)} = 0$$

$$(ii) \quad \phi'(0) = \lim_{s \rightarrow \infty} s^2 \bar{f}(s)$$

Now

$$\bar{f}(s) = \left(\frac{\rho}{\rho+s}\right)^a \left(\omega_1 \frac{\rho_1}{\rho_1+s} + \omega_2 \frac{\rho_2}{\rho_2+s}\right)$$

$$\phi'(0) = \begin{cases} \omega_1 \rho_1 \rho + \omega_2 \rho_2 \rho & \text{if } a = 1 \\ 0 & \text{if } a > 1 \end{cases}$$

(iii) $\phi(x)$ as $x \rightarrow \infty$

(a) $\rho = \min(\rho, \rho_1, \rho_2)$, $\rho \neq \rho_1 \neq \rho_2$

$$\lim_{x \rightarrow \infty} \phi(x) = \lim_{x \rightarrow \infty} \frac{f(x)/x^{a-1}e^{-\rho x}}{S(x)/x^{a-1}e^{-\rho x}}$$

$$= \frac{\omega_1 \rho_1 \left(\frac{\rho}{\rho-\rho_1}\right)^a \left[-\frac{(\rho-\rho_1)^{a-1}}{(a-1)!} \right] + \omega_2 \rho_2 \left(\frac{\rho}{\rho-\rho_2}\right)^a \left[-\frac{(\rho-\rho_2)^{a-1}}{(a-1)!} \right]}{\frac{\rho^{a-1}}{(a-1)!} + \omega_1 \left(\frac{\rho}{\rho-\rho_1}\right)^a \left[-\frac{(\rho-\rho_1)^{a-1}}{(a-1)!} \right] + \omega_2 \left(\frac{\rho}{\rho-\rho_2}\right)^a \left[-\frac{(\rho-\rho_2)^{a-1}}{(a-1)!} \right]}$$

$$= \frac{\rho \rho_1 \rho_2 - \rho^2 (\omega_1 \rho_1 + \omega_2 \rho_2)}{\rho_1 \rho_2 - \rho (\omega_1 \rho_1 + \omega_2 \rho_2)}$$

$$= \rho$$

(b) $\rho_1 = \min(\rho, \rho_1, \rho_2)$, $\rho \neq \rho_1 \neq \rho_2$

$$\lim_{x \rightarrow \infty} \phi(x) = \frac{f(x)/e^{-\rho_1 x}}{S(x)/e^{-\rho_1 x}}$$

$$= \frac{\omega_1 \rho_1 \left(\frac{\rho}{\rho-\rho_1}\right)^a}{\omega_1 \left(\frac{\rho}{\rho-\rho_1}\right)^a}$$

$$= \rho_1$$

(c) $\rho_2 = \min(\rho, \rho_1, \rho_2)$, $\rho \neq \rho_1 \neq \rho_2$

$$\lim_{x \rightarrow \infty} \phi(x) = \rho_2$$

(d) $\rho = \rho_1 = \rho_2$, the combination becomes a Special Erlangian distribution having

$$\lim_{x \rightarrow \infty} \phi(x) = \rho$$

(e) $\rho_1 = \rho_2 < \rho$ then

$$\lim_{x \rightarrow \infty} \phi(x) = \frac{\omega_1 \rho_1 \left(\frac{\rho}{\rho-\rho_1}\right)^a + \omega_2 \rho_2 \left(\frac{\rho}{\rho-\rho_2}\right)^a}{\omega_1 \left(\frac{\rho}{\rho-\rho_1}\right)^a + \omega_2 \left(\frac{\rho}{\rho-\rho_2}\right)^a}$$

$$= \omega_1 \rho_1 + \omega_2 \rho_2$$

$$= \rho_1 = \rho_2$$

(f) $\rho = \rho_1 < \rho_2$

Then

$$\lim_{x \rightarrow \infty} \phi(x) = \rho$$

It can be concluded from above that

$$\lim_{x \rightarrow \infty} \phi(x) = \min(\rho, \rho_1, \rho_2)$$

Appendix V

Moments of Stage Combinations

Series of Identical Stages

The Laplace transform of the probability density function is

$$\bar{f}(s) = \left(\frac{\rho}{s+\rho}\right)^a$$

Differentiating successively and substituting $s = 0$

$$\bar{f}^{(r)}(0) = (-1)^r \frac{1}{\rho^r} \prod_{k=1}^r (a+k-1)$$

The r th moment is therefore

$$m_r = \frac{1}{\rho^r} \prod_{k=1}^r (a+k-1)$$

Two Series Stages in Parallel

The r th moment in this case can be written as

$$m_r = \frac{\omega_1}{\rho_1} \prod_{k=1}^r (a_1 + k - 1) + \frac{\omega_2}{\rho_2} \prod_{k=1}^r (a_2 + k - 1)$$

Series stages in series with a distinctive stage

$$\bar{f}(s) = \left(\frac{\rho}{s + \rho} \right)^a \left(\frac{1}{s + \rho_1} \right)$$

That is

$$(s + \rho)^a (s + \rho_1) \bar{f}(s) = \rho_1 \rho^a$$

Differentiating both sides

$$(s + \rho)(s + \rho_1) \bar{f}'(s) + \{a(s + \rho) + (s + \rho)\} \bar{f}(s) = 0 \quad (1)$$

Differentiating once again

$$(s + \rho)(s + \rho_1) \bar{f}''(s) + [2(s + \rho) + (a + 1)(s + \rho_1)] \bar{f}'(s) + (a + 1) \bar{f}(s) = 0 \quad (2)$$

Differentiating r times

$$(s + \rho)(s + \rho_1) \bar{f}^{(r)}(s) + \{r(s + \rho) + (a + r - 1)(s + \rho_1)\} \bar{f}^{(r-1)}(s) + (r - 1)(a + r - 1) \bar{f}^{(r-2)}(s) = 0 \quad (3)$$

Putting

$$s = 0$$

and

$$\bar{f}^{(r)}(0) = (-1)^r m_r$$

From Equation (1)

$$\rho \rho_1 m_1 = a \rho_1 + \rho \quad \text{for } r = 1$$

From Equation (2)

$$\rho \rho_1 m_2 - \{2\rho + (a + 1)\rho_1\} m_1 = -(a + 1) \quad \text{for } r = 2$$

From Equation (3)

$$\rho \rho_1 m_r - \{r\rho + (a + r - 1)\rho_1\} m_{r-1} + (r - 1)(a + r - 1) m_{r-2} = 0 \quad \text{for } r > 2$$

In the matrix form

$$[A] [m] = [B] \quad (4)$$

where

$$A = [a_{ij}] \text{ is the coefficient matrix such that}$$

$$a_{ij} = 0 \quad \text{if } j > i \quad \text{or } j < i - 2$$

$$a_{ij} = \rho \rho_1 \quad \text{if } i = j$$

$$a_{ij} = -\{i\rho + (a + i - 1)\rho_1\} \quad \text{if } j = i - 1$$

and

$$a_{ij} = (i - 1)(a + i - 1) \quad \text{if } j = i - 2.$$

$$m = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_r \end{bmatrix}$$

$$B = \begin{bmatrix} \rho + a\rho_1 \\ -(a + 1) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The r moments can be found by solving the set of linear Equations (4).

Series stages in series with two parallel stages

This is equivalent to two 'series stages in series with a distinctive stage' in parallel. The r th moment for the whole combination is obtained by

$$[m] = [m_1] + [m_2]$$

$[m_1]$ and $[m_2]$ are found by equations,

$$[A_1] [m_1] = [B_1]$$

and

$$[A_2] [m_2] = [B_2]$$

In this case

$$[B_1] = \begin{bmatrix} (\rho + a\rho_1)\omega_1 \\ -(a+1)\omega_1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

and

$$[B_2] = \begin{bmatrix} (\rho + a\rho_2)\omega_2 \\ -(a+1)\omega_2 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

Appendix VI

Calculation of the Jacobian Matrix for two series stages in parallel

Assuming a_1 and a_2 , the remaining three parameters ρ_1 , ρ_2 and ω_1 can be calculated by matching the first three moments. Therefore

$$X_0 = [\rho_{10}\rho_{20}\omega_{10}]^t, \quad \phi = [\phi_1\phi_2\phi_3]^t$$

The elements of the j th column of the Jacobian matrix can be obtained by differentiating

$$\phi_j = \frac{\omega_1}{\rho_1^j} \prod_{k=1}^j (a_1 + k - 1) + \frac{\omega_2}{\rho_2^j} \prod_{k=1}^j (a_2 + k - 1) - M_j$$

That is

$$\frac{\partial \phi_j}{\partial \rho_1} = -\frac{j\omega_1}{\rho_1^{j+1}} \prod_{k=1}^j (a_1 + k - 1)$$

$$\frac{\partial \phi_j}{\partial \rho_2} = -\frac{j\omega_2}{\rho_2^{j+1}} \prod_{k=1}^j (a_2 + k - 1)$$

and

$$\frac{\partial \phi_j}{\partial \omega_1} = \frac{1}{\rho_1^j} \prod_{k=1}^j (a_1 + k - 1) - \frac{1}{\rho_2^j} \prod_{k=1}^j (a_2 + k - 1)$$

Series Stages in Series with Two Parallel Stages

Assuming the number of stages to be a , the remaining four parameters, ρ_1 , ρ_2 , ρ and ω_1 can be calculated by matching the first four moments. The vector

$$\phi = m - M$$

where m and M are vectors of the stage model moments and the moments of the distribution to be approximated. Since $m = m_a + m_b$, the Jacobian of ϕ at X_0 becomes

$$\phi'(X_0) = m'_a(X_0) + m'_b(X_0)$$

$m'_a(X_0)$ and $m'_b(X_0)$ can be obtained by differentiating and solving

$$A_1 m_a = B_a \quad \text{and} \quad A_2 m_b = B_b$$

This solution can be obtained using Gauss elimination.

Index

- Absorbing state, 44
- Age specific failure rate, 11-12
- Algorithm for minimal cut sets, 127, 129
- Allocation, 4
- Alternate designs, 6
- Alternating renewal process, 80
- Analytical modelling, 5
- Antithetic variables, 224
- Approximation for non-Markovian model, 198
- Asynchronous method, 214
- Availability,
 - steady state, 3, 221
 - time specific, 2, 69, 220
- Average values, 87
- Block schematic diagram, 109
- Boole's inequality, 118, 119
- Box and Muller method, 219
- Central limit theorem, 223
- Central moment, 16
- Change out time, 173, 179
- Chapman-Kolmogorov equation, 35, 49
- Characteristic function, 17-18
- Closed communicating class, 39
- Coefficient of excess, 17
- Coefficient of skewness, 17
- Complex transition rates, 207-10
- Conditions of mergeability, 135-41
- Confidence interval, 222-3
- Connected sub network, 116
- Continuous parameter Markov chains,
 - 48-61, 177
 - dishonest process, 49
 - equilibrium probability distribution, 53
 - first passage time, 57
 - forward equations, 51
 - honest process, 49
 - Kolmogorov differential equation, 52
 - mean cycle time, 61
 - mean duration, 60
 - mean time to first failure, 57, 59
 - time homogenous, 49
 - transient behaviour, 52
 - transition rate, 51
- Continuous random variable, 8-9
- Continuous state, 170
- Control variates, 224
- Correlation coefficient, 15
- Countable set, 7
- Countably infinite, 7
- Covariance, 15
- Cramer's rule, 227
- Cut set, 116
- Cut set approach for frequency, 120-5
- Cut set manipulation of probabilities, 119
- Cut set methods, 119
- Cyclic set, 39
- Data collection, 5
- Decomposition approach, 91, 106-7
- Denumerable set, 7
- Dependent failures, 91
- Device of stages, 179-195
 - derivation of characteristics, 181
 - determination of parameters, 180-91
 - selection of stage combinations, 179-80
- Diagonal matrix, 45
- Discrete random variable, 7-8
- Dishonest process, 49
- Eigenvalue, 41-3, 71
- Eigenvectors, 41-3, 71
- Equilibrium alternating renewal process, 80
- Equilibrium distribution, 39-40, 53

- Equivalent transition rate, 87, 135
- Ergodic set, 39
- Estimation of reliability measures, 220-21
- Events, 8
- Event space, 91
- Expectation, 13
- Exponential distribution, 22-5
 - distribution of residual life time, 24
 - random sampling, 208
- Factors of influence method, 4
- Failure modes and effects analysis, 90, 109, 213
- Failure modes, effects and criticality analysis, 5, 90
- Fatigue failures, 31
- Final value theorem, 21
- Finite set, 7
- First passage time, 44-5, 57
- Fixed interval method, 214
- Fluctuating environment, 144-5
- Force of mortality, 11-12
- Forward equation, 51, 68
- Fractional duration, 2, 69, 71-2, 221
- Frequency
 - interval, 70, 72
 - steady state, 3, 74
 - time specific, 66, 67
- Frequency balance, 96
- Frequency balance equation, 101
- Frequency balancing approach or technique, 48, 63-79, 225
- Frequency equilibrium, independent components, 78-9
- Functional diagram, 89
- Fundamental matrix, 45
- Gamma distribution, 31, 183
- Gamma function, 30
- Gauss elimination, 133
- Gauss-Jordan method, 133, 227
- Hazard function, 11-12
- Honest process, 49
- Incomplete gamma distribution, 32
- Independent components, 206-7
- Initial moment, 16
- Initial value theorem, 21
- Interconnected power systems, 106
- Interstate transition rate, 63-6
- Interval frequency, 2, 70, 72, 221
- Inverse of probability distribution method, 216-17
- Jacobian matrix, 194, 242-3
- Key component, 106
- Kolmogorov differential equation, 52
- Laplace transform, 20
- Large systems, problem areas, 132-4
- Law of large numbers, 13
- Logic diagram, 109
- Lognormal distribution, 28-9, 195
- Maintainable systems, 89
- Markov chains, 36-48, 177
 - absorbing state, 44
 - closed communicating class, 39
 - cyclic or periodic, 39
 - diagonal matrix, 45
 - equilibrium distribution, 39-40
 - ergodic set, 39
 - first passage time, 44-5
 - fundamental matrix, 45
 - mean time to first failure, 44, 57, 59
 - steady state probabilities, 40
 - stochastic matrix, 37
 - time specific behaviour, 41
 - transient set, 39
 - transition matrix, 37
- Mathematical model, 211, 212
- Mean cycle time, 3, 61, 75, 84, 221
- Mean duration, 3, 60, 75
- Mean first passage time, 3
- Mean passage time, 3
- Mergeability, 88
 - conditions of, 135-41
- Method of supplementary variables, 165-76
- Mixture of Erlangian distributions, 187
 - hazard function of, 230-3
 - moments of, 239-40
- Modified renewal process, 85
- Moments, 16-17
- Moment generating function, 18-20
- Moments of stage combinations, 239-41

- Monte Carlo technique, 222
- Multiplicative congruential method, 212
- Network method (approach), 91, 109-29
- Network reduction procedure, 110-11
- Newton-Raphson method, 192, 194-5
- Next event method, 214
- Non-maintained systems, 89
- Non-Markovian systems, 164
- Non-singular coefficient matrix, 227
- Normal distribution, 25-28
 - random sampling of, 219-20
- Operational methods, 19-22
- Parallel systems, 103-4
 - m/n parallel, 104-5
 - parallel redundant, 103
- Path, 116
- Physical diagram, 109
- Poisson distribution, 24
- Possibility space, 7
- Probability density function, 10-11
- Probability distributions, 22-33
- Probability generating function, 20
- Probability laws, 9-12
- Probability mass function, 10
- Proportionate allocation technique, 215-16
- Pseudo random numbers, 213
- Pumping system, 7
- Random number generation, 212
- Random sampling, 214-20
- Random variables, 8-9
- Reliability, 2
- Reliability block diagram, 109
- Reliability, planning, 1
- Reliability program, 1
- Reliability program elements, 1
- Reliability report, 6
- Renewal density, 84-5
- Renewal theory, 80
- Residual life time, 81
- Sample space or sample description space, 7
- Semi-Markov process, 176-9
- Sensitivity analysis, 211
- Sequential truncation, 154, 158
- Series stages in series with distinctive stage, 189-91
 - hazard function of, 232-4
 - moments of, 240-1
- Series stages in series with two parallel stages, 191, 234-9
 - moments of, 241-2
- Series system, 97-102
 - dependent case, 99-100
 - independent case, 97-8
 - with spare, 100-102
- Similar familiar technique, 4
- Simple path, 116
- Simpson's rule, 199
- Simulation, 5, 211
- Simulation model, 213
- Simultaneous linear equations, 227-9
- Skew coefficient, 17
- Special distributions, 22-33
- Special Erlangian distribution, 32, 183
 - moments of, 239
 - random sampling of, 218-19
- Stages in parallel, 185-8
- Stages in series, 182-5
- Standard deviation, 14
- Standard normal distribution, 26
- State space, 91
- State space approach, 91
- State space truncation, 151-3
- State transition diagram, 91
- Stochastic matrix, 37
- Stochastic processes, 33-6
 - indexing parameter, 33
 - parameter space, 33
 - probability distribution, 35-6
 - realization of, 34
- Stratification, 224
- Survivor function, 11
- Synchronous timing method, 214
- System, 89
 - description of, 89-90
- System reliability, 2
 - quantitative measures, 2
 - reference indices, 2
- t -distribution, 222
- Testing, 5
- Tie set, 117
- Tie set manipulation of probabilities, 117-18
- Tie set methods, 116

248 *System Reliability Modelling and Evaluation*

- Time homogenous Markov process, 49
 - hazard function of, 230-3
 - moments of, 239-40
- Time specific availability, 2
- Time specific behaviour of Markov chains, 41
- Time specific frequency, 66-7
- Timing controls, 214
- Transform methods, 17-22
- Transient behaviour, 52
- Transient set, 39
- Transition matrix, 37
- Transition rates, 51, 63-6
- Triangular distribution function, 217
- Truncated normal distribution, 27
- Two series stages in parallel, 187
 - Two state Markov process, 54
- Uncountable or uncountably infinite, 8
- Uniform probability density function, 217
- Variance, 14
- Variance reducing techniques, 223-4
- Vertex state, 120
- Waiting time, 177
- Weibull distribution, 30-1